

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

AJ

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C12Q 1/68, C07K 14/705, C12N 9/64, C07K 14/72, C12N 9/02	A2	(11) International Publication Number: WO 00/22166 (43) International Publication Date: 20 April 2000 (20.04.00)
(21) International Application Number: PCT/IB99/01678 (22) International Filing Date: 13 October 1999 (13.10.99) (30) Priority Data: 60/104,286 14 October 1998 (14.10.98) US 60/104,302 14 October 1998 (14.10.98) US (71) Applicant (for all designated States except US): EURONA MEDICAL AB [SE/SE]; Kungsgangsvägen 29, S-751 06 Uppsala (SE). (72) Inventors; and (75) Inventors/Applicants (for US only): NORBERG, Leif, Tor- bjorn [SE/SE]; Backvägen 14B, S-756 52 Uppsala (SE). ANDERSSON, Maria, Kristina [SE/SE]; Backvägen 14B, S-756 52 Uppsala (SE). LINDSTROM, Per, Harry, Rut- ger [SE/SE]; Svankarrevägen 26A, S-756 33 Uppsala (SE). JONSSON, Lena [SE/SE]; Gustaf Kjellbergsvag 20, S-756 42 Uppsala (SE).		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>Without international search report and to be republished upon receipt of that report.</i>
(54) Title: GENES FOR ASSESSING CARDIOVASCULAR STATUS AND COMPOSITIONS FOR USE THEREOF (57) Abstract The present invention provides methods for assessing cardiovascular status in an individual, which comprise determining the sequence at one or more polymorphic positions within the human genes encoding ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin receptor, and β -adrenoceptor. The invention also provides isolated nucleic acids encoding polymorphisms in genes encoding ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin receptor and β -adrenoceptor, nucleic acid probes that hybridized to polymorphic positions, kits for the prediction of cardiovascular status, and nucleic acid and peptide targets for use in identifying candidate cardiovascular drugs.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

**GENES FOR ASSESSING CARDIOVASCULAR STATUS
AND COMPOSITIONS FOR USE THEREOF**

FIELD OF THE INVENTION

5 The present invention relates to genetic polymorphisms and polymorphism patterns useful for assessing cardiovascular status in humans. More particularly, the invention relates to identifying and using polymorphism patterns comprising a polymorphism in the a gene encoding a polypeptide selected from the group consisting of ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin
10 receptor, and β -adrenoceptor to predict a treatment outcome or likelihood of developing cardiovascular disease, and to assist in diagnosis and in prescription of an effective therapeutic regimen.

BACKGROUND OF THE INVENTION

15 The renin-angiotensin-aldosterone system (RAAS) plays an important role in cardiovascular physiology in mammals. Specifically, RAAS regulates salt-water homeostasis and the maintenance of vascular tone. Stimulation or inhibition of this system raises or lowers blood pressure, respectively, and disturbances in this system may be involved in the etiology of, for example, hypertension, stroke, and myocardial
20 infarction. The RAAS system may also have other functions such as, *e.g.*, control of cell growth. The renin-angiotensin system includes renin, angiotensin converting enzyme

(ACE), angiotensinogen (AGT), type 1 angiotensin II receptor (AT1), type 2 angiotensin II receptor (AT2) and aldosterone synthase.

International Patent Application No. PCT/IB98/00475, filed April 1, 1998, discloses for the first time an association of polymorphism patterns in ACE, AGT, and
5 AT1 genes with cardiovascular status, particularly with the ability to predict the therapeutic outcome of a particular treatment regimen.

Other genes express polypeptides involved in pathways other than RAAS that also play a role in the regulation of cardiovascular physiology, although it is believed that prior to the present invention, no known association of polymorphism patterns in
10 these genes with cardiovascular status has been observed or reported. Among these other regulators of cardiovascular physiology are endothelin, endothelin receptor, and β -adrenergic receptors 1 and 2.

RAAS Pathway Components

AGT is the specific substrate of renin, an aspartyl protease. The human
15 AGT gene contains five exons and four introns which span 13Kb (Gaillard *et al.*, DNA, 1989, 8:87-99; Fukamizu *et al.*, J. Biol. Chem., 1990, 265:7576-7582). The first exon (37 bp) codes for the 5' untranslated region of the mRNA. The second exon codes for the signal peptide and the first 252 amino acids of the mature protein. Exons 3 and 4 are shorter and code for 90 and 48 amino acids, respectively. Exon 5 contains a short coding
20 sequence (62 amino acids) and the 3'-untranslated region.

Plasma AGT is synthesized primarily in the liver and its expression is positively regulated by estrogens, glucocorticoids, thyroid hormones, and angiotensin II (Ang II) (Clauser *et al.*, Am. J. Hypertension, 1989, 2:403-410). Cleavage of the amino-terminal segment of AGT by renin releases a decapeptide prohormone, angiotensin-I,
25 which is further processed to the active octapeptide angiotensin II by the dipeptidyl carboxypeptidase designated angiotensin-converting enzyme (ACE). Cleavage of AGT by renin is the rate-limiting step in the activation of the renin-angiotensin system.

Several epidemiological observations indicate a possible role of AGT in blood pressure regulation. A highly significant correlation between plasma AGT
30 concentration and blood pressure has been observed in epidemiological studies (Walker *et al.*, J. Hypertension, 1979, 1:287-291). Interestingly, a number of allelic dimorphisms have been identified in the AGT gene. The frequency of at least two of them (174M and 235T) have been partially characterized and in certain populations shown to be

significantly elevated in hypertensive subjects (Jeunemaitre *et al.*, Cell, 1992, 71:169-180). In addition, a specific polymorphism, 235T, has been suggested to be directly involved in coronary atherosclerosis (Ishigami *et al.*, Circulation, 1995, 91:951-4). Furthermore, the presence of A or G at position 1218 in the AGT regulatory region has
5 been correlated with differences in *in vitro* transcriptional capacity for this gene (Inoue *et al.*, J. Clin. Invest., 1997, 99:1786). However, the foregoing are studies involving only one or at most two polymorphisms. Furthermore, the sole disclosed use is susceptibility to disease.

The human ACE gene is also a candidate as a marker for hypertension and
10 myocardial infarction. ACE inhibitors constitute an important and effective therapeutic approach in the control of human hypertension (Sassahō *et al.*, Am. J. Med., 1987, 83:227-235). In plasma and on the surface of endothelial cells, ACE converts the inactive angiotensin I molecule (Ang I) into active angiotensin II (Ang II) (Bottari *et al.*, Front. Neuroendocrinology, 1993, 14:123-171). Another ACE substrate is bradykinin, a potent
15 vasodilator and inhibitor of smooth muscle cell proliferation, which is inactivated by ACE (Ehlers *et al.*, Biochemistry, 1989, 28:5311-5318; Erdos, E.G., Hypertension, 1990, 16:363-370; Johnston, C.I. *Drugs* (suppl. 1), 1990, 39:21-31).

Levels of ACE are very stable within individuals, but differ greatly between individuals. A greater risk of myocardial infarction has been identified in a group of
20 subjects with an ACE polymorphism designated ACE-DD (Cambien *et al.*, Nature, 1992, 359:641-644), and a 12-fold greater risk of myocardial infarction has been identified in a subgroup of patients having a combination of the ACE polymorphism ACE-DD and one of the AGT polymorphisms (235T) described above (Kamitani *et al.*, Hypertension, 1994, 24:381). Recently, six ACE polymorphisms were identified and characterized (Villard *et al.*, Am. J. Human Genet., 1996, 58:1268-1278).
25

The vasoconstrictive, cell growth-promoting and salt conserving actions of angiotensin II are mediated through binding to and activation of angiotensin receptors, of which at least two types have been cloned (AT1 and AT2). The type 1 Ang II receptor (AT1), a G-protein-coupled seven transmembrane domain protein, is widely distributed in
30 the body and mediates almost all known Ang II effects (Fyhrquist *et al.*, J. Hum. Hypertension, 1995, 5:519-524).

Several polymorphisms have been identified in the AT1 receptor gene. Initial studies suggest that at least one of them is more frequent in hypertensive subjects

(AT¹¹⁶⁶C)(Bonnardeaux *et al.*, Hypertension, 1994, 24:63-69). This polymorphism, combined with the ACE-DD polymorphism, has been shown to correlate strongly with the risk of myocardial infarction (Tiret *et al.*, Lancet, 1994, 344:910-913).

5 **Endothelin (ET) Regulation of Cardiovascular Physiology**

 Endothelin is a potent vasoconstrictive peptide characterized by long lasting action. It was first discovered as a vasoconstricting factor in conditioned medium (Hickey *et al.*, Am. J. Physiol., 1985, 248:C550), and subsequently purified and characterized (Yanagisawa *et al.*, Nature, 1988, 332:411). ET is produced as
10 preproendothelin, which is cleaved after removal of the signal sequence by an endopeptidase, followed by cleavage with endothelin converting enzyme (Xu *et al.*, Cell, 1994, 78:473; Shimada *et al.*, J. Biol. Chem., 1994, 269:18274). Analysis of the human ET gene has revealed the existence of two additional ET-like peptides expressed in various tissues, termed ET-2 and ET-3 (Inoue *et al.*, Proc. Natl. Acad. Sci. USA, 1989, 86:2863).
15 The first endothelin was accordingly termed ET-1.

 One of the important discoveries following characterization of ETs was the discovery of two ET receptors, ET_A and ET_B (Arai *et al.*, Nature, 1990, 348:730; Sakumi *et al.*, Nature, 1990, 348:782). Both belong to the family of heptahelical G-protein coupled receptors. There is 68% amino acid identity between the two receptor subtypes.
20 ET_A exists as a single copy gene located on human chromosome 4 (Hosoda *et al.*, J. Biol. Chem., 1992, 267:18797; Cyr *et al.*, Biochem. Biophys. Res. Commun., 1991, 181:184). ET_B exists as a single copy gene located on human chromosome 13 (Arai *et al.*, J. Biol. Chem., 1993, 268:3463-70), although a splice variant of ET_B has been found (Shyamala *et al.*, Cell. Mol. Biol. Res., 1994, 40:285-96). The cDNA sequence of ET_A has been
25 deposited with GenBank with accession number S57498.

 In the early stage of ET research, a great number of pharmacological studies suggested that the responses to ETs could be divided into two groups according to the pharmacological potency of the three peptides. Indeed, these two receptors, ET_A and ET_B, are distinct in their ligand binding affinity and distribution in tissues and cells. ET_A
30 has a high affinity to ET-1 and ET-2, but a low affinity to ET-3. ET_B has equally potent affinities to all three endogenous ETs. ET_A exists on smooth muscle and mediates vasoconstriction. In contrast, ET_B exists on endothelium and mediates the release of

relaxing factors such as nitric oxide and prostacycline. However, several reports demonstrated that ET_B on some vascular smooth muscle also mediated vasoconstriction.

Cloning of the *ET receptor* gene facilitated the development of ET-receptor antagonists, such as BE-18257B, and BQ-123 and FR139317, two derivatives of
5 BE-18257B (See Masaki, Cardiovascular Res., 1998, 39:530). Many selective and non-selective antagonists for ET_A and ET_B have emerged.

Although the pathophysiological role of ET is still unclear, ET antagonists demonstrated significant beneficial effects in pathological conditions, including congestive heart failure, pulmonary hypertension, cerebrovascular spasm after subarachnoid
10 hemorrhage, acute renal failure, and essential hypertension. ET or ET receptor knockout mice have also provided important information regarding the physiological and pathophysiological significance of ET (Masaki, *supra*). In particular, mice with a knockout of *ET-3* or the *ET_B receptor* genes exhibit phenotypic changes that resemble Hirschsprung's disease, a human hereditary syndrome associated with a missense mutation
15 of the *ET_B* gene (Pfaffenberger, *et al.*, Cell, 1994, 79:1257).

Despite advances in understanding the role of the endothelin pathway in the treatment of cardiovascular diseases, questions remain. Not all experimental models of hypertension respond to endothelin antagonists, and it remains uncertain whether endothelin antagonists improve cardiac structure and function beyond the benefits of blood
20 pressure reduction (Moreau, Cardiovascular Res., 1998, 39:534). Thus, there is a need in the art for a reliable and effective means for predicting whether endothelin antagonists will be effective for treating hypertension in a given individual.

β-Adrenergic Receptors (β-Adrenoceptors)

25 The adrenoceptors fall into three major groups, α₁, α₂, β, within each of which further subtypes can be distinguished pharmacologically (Lüllmann, *et al.* in Color Atlas of Pharmacology, New York, 1993). Adrenergic receptors are all G-protein linked. They are involved in regulation of the cardiovascular system, and in the control of metabolic activity, *e.g.*, insulin secretion and glucose release. They also mediate
30 constriction or relaxation of smooth muscle cells in the respiratory, gastrointestinal, and genitourinary tracts (Berne and Levy, Principles of Physiology (2nd Ed.), Mosby-Year Books, Inc., 1996, pp. 691-696).

Adrenoceptors are targets for epinephrine and norepinephrine, which are representatives of the family of monoamine neurotransmitters. Epinephrine has equally high affinity for all α - and β -receptors while norepinephrine differs from epinephrine by its low affinity for β_2 -receptors (The Biochemical Basis of Neuropharmacology, (7th Ed.) New York, 1996, pp. 226-292). The adrenoceptors themselves interact preferentially with three different classes of G-proteins: G_s (β -adrenoceptors) mediating activation of adenylate cyclase, G_i (α_2 -adrenoceptors) mediating inhibition of adenylate cyclase, and G_q (α_1 -adrenoceptors) mediating activation of phospholipase C (Hieble *et al.*, J. of Med. Chem., 1995, 38:3415-3444).

10 The pharmacological interest of adrenoceptors is mainly for the treatment of cardiovascular diseases, *e.g.* through the development of β -antagonists, α_1 -antagonists and α_2 -agonists to treat hypertension, but they are also considered important for the treatment of asthma (β_2 -agonists).

 The β_2 -adrenergic receptor is expressed on a number of cell types, *e.g.*,
15 bronchial smooth muscle, where its activation results in relaxation and bronchial dilatation. These receptors are also being expressed on epithelial cells, vascular endothelium, alveolar walls, immune cells, and presynaptic nerve terminals (Liggett, Am. J. Respir. Crit. Care. Med., 1997, 156:S156-S162). Cardiac cells express mainly β_1 -, but also a small fraction of β_2 -adrenoceptors (Collins *et al.*, Biochimica et Biophysica Acta,
20 1993, 1172:171-174). β_1 -adrenoceptors are also expressed in brain and pineal gland.

β -Adrenoceptor Function

 The β_1 - and β_2 -adrenergic receptors are coupled to a G_s -protein complex, which activates adenylate cyclase. Agonist binding to the β_1 -receptor, located in cardiac
25 muscle cells, mediates increased contractility and cardiac output (Lüllmann, *supra*). Agonist binding to the β_2 -receptor, located in peripheral vascular arteries, mediates vasodilation by increasing the amount of cAMP, and thereby inhibiting activation of myosin kinase, which is necessary for smooth muscle cell constriction (*ibid.*) Activation of cAMP in cardiac cells by agonist binding to both β_1 - and β_2 -adrenoceptors activates the
30 cAMP-dependent protein kinase (PKA) (Castellano and Böhm, Hypertension, 1997, 29:715-722). β -adrenoceptors also regulate the control of melatonin production in the pineal gland, by the cAMP activation of one of the enzymes (5-HT-N-acetyl transferase) involved in the synthesis of melatonin (Collins *et al.*, *supra*).

The β_2 -receptors mediate increased conversion of glycogen to glucose (glycogenolysis) in both the liver and skeletal muscle (H. Lüllmann, *et al.*, *supra*), and stimulate influx of potassium into muscle cells to prevent hyperkalemia (Berne and Levy, *supra*).

5 The β -receptors are regulated on the protein level by desensitization. The initial desensitization process results from the phosphorylation of serine and threonine residues in the cytoplasmic tail or third intracellular loop by several protein kinases, including β ARK and PKA (Hieble *et al.*, *supra*). β ARK phosphorylates specific serine or threonine residues in the C-terminal of receptors that are occupied by an agonist. The
10 phosphorylation triggers binding of the cytosolic protein β -arrestin and results in the uncoupling from G_{sa} . PKA is activated by cAMP and phosphorylates the β_2 -adrenoceptor by a relatively slow process. The phosphorylated receptor loses the ability to activate G_s (Castellano and Böhm, *supra*). Prolonged interaction of agonists with adrenoceptors generally results in receptor desensitization.

15

β -Adrenoceptor Gene Structure

The human β_1 -adrenoceptor gene is located on the long arm of chromosome 10, the same chromosome as for the α_{2A} -adrenoceptor gene. The coding sequence of this gene is deposited with GenBank, accession number X69168. The regulatory region is also
20 deposited with GenBank, accession number J03019. It codes for an intronless gene product of 1431 base pairs (Hall, Thorax, 1996, 51:351-353). Both the promoter and the coding region of the gene are rich in G and C residues, which make up greater than 70% of the bases. The promoter does not contain any paired consensus TATA box and CAAT box elements but instead clusters with an inverted CAAT box and SP₁ or AP-2 binding
25 motifs. This type of receptor, reminiscent of "housekeeping genes", has been described for other G-protein coupled receptors as well (Collins *et al.*, *supra*).

The human β_2 -adrenoceptor gene is located on the long arm of chromosome 5, the same chromosome as the α_{1B} -adrenoceptor gene. The coding sequence has been deposited with GenBank, with accession numbers M15169, J02728, or M16106. The
30 regulatory region sequence is also deposited with GenBank, accession number Y00106. It codes for an intronless gene product of 1239 base pairs (Hall, *supra*). The promoter region is 200-300 bases 5' of the translation initiation codon, and it can form strong secondary structures due to high G-C content. There are two TATA boxes (separated by roughly 10

bp) and a CAAT box located approximately 30 and 80 base pairs upstream, respectively, from the mRNA start region.

β -Adrenoceptor Gene Regulation

5 There are some regulatory regions identified in the promoter region of the β_1 -adrenoceptor gene: a cAMP response element (CRE), a consensus thyroid response element (TRE), and a glucocorticoid response element (GRE). This is consistent with the evidence that both thyroid hormone and corticosteroids affect adrenergic sensitivity in both heart and adipose tissue. The CRE region might have a self-regulatory function, as
10 has been shown for the β_2 -adrenoceptor gene (Collins et al., *supra*).

 There are several regulatory domains in the 5' flanking region of the β_2 -adrenoceptor. Among these is a cAMP-responsive element (CRE), which is recognized and stimulated by a phosphoprotein called CRE binding protein (CREB). CREB is partially under the control of PKA-dependent phosphorylation processes. This is seen as
15 an increase in β_2 -adrenoceptor mRNA level in the early phase after exposure to β -agonists. However, the level of mRNA is decreased after prolonged exposure to agonists, probably mediated by a shortening of mRNA half-life (Castellano and Böhm, *supra*). It has also been shown that transcription of the β_2 -adrenoceptor gene is upregulated by stimulation with glucocorticoids in a variety of tissues (Collins et al., *supra*). In the 3' flanking region
20 there are sequences homologous to glucocorticoid response elements. These might be responsible for the increased expression of β_2 adrenoceptor observed in transfected cells after treatment with hydrocortisone (Emorine and Marullo, Proc. Natl. Acad. Sci. USA, 1987, 84:6995-6999).

β -Adrenoceptor Protein Structure

25 The proposed model for β -adrenoceptors is like most of the G-protein binding receptors, a seven α -helical transmembrane structure, where the seven α -helices are radially arranged around a central "pore", in which the receptor ligands bind. The β -adrenoceptors have an extracellular glycosylated N-terminus, and an intracellular
30 C-terminus. The β_1 -receptor consists of 477 amino acids; the β_2 -receptor consists of 413 amino acids.

 The overall amino acid identity of human β_1 - and β_2 - adrenoceptors is only 54%. However, it is likely that the pharmacological differences between β_2 -receptors and

β_1 -receptors are due to subtle changes in orientation of the primary binding sites, resulting in a slightly different binding site rather than to specific amino acid substitutions (Hieble *et al.*, J. Med. Chem. 1995, 38:3415-3444).

Site-directed mutagenesis has demonstrated that an aspartic acid residue,
5 Asp-113, located in the third transmembrane-spanning helix, and two serine residues, Ser-204 and Ser-207, are required for full agonist binding to the β_2 -adrenoceptor. The β_1 -adrenoceptor contains identical amino acid residues located in corresponding positions to those shown to be important for agonist binding to the β_2 -adrenoceptor. Another
10 aspartic acid residue, Asp-79, located in the second α -helix of both β -receptors is highly conserved in G-protein coupled receptors (Hieble *et al.*, *supra*). Ser-319 has a potential role in agonist binding to the β_2 -adrenoceptor.

Mutation of Tyr-350, located in the cytoplasmic tail of the β_2 -receptor, interferes with coupling of the receptor to G_s (Hieble *et al.*, *supra*). Also, palmitoylation
15 of Cys-341 in the C-terminal enables the β_2 -adrenoceptor to form a fourth intracytoplasmic loop, which increases the ability of the agonist-bound receptor to mediate adenylyl cyclase stimulation (Strosberg, Preotein Science, 1993, 2:1198-1209).

β -Adrenoceptors as Drug Targets

No cause of disease can be identified in 80-90% of patients with
20 hypertension. They have so-called essential hypertension, which affects 5-10% of the general population, and is the most common cause of disease in developed countries (J. Axford, *Medicine*, Blackwell Science Ltd., 1996, 10.119-10.130). Betablockers have been widely used in the treatment of hypertension. They are particularly useful for the treatment of juvenile hypertension with tachycardia and high cardiac output. Betablockers
25 or beta-adrenergic blockers were first introduced as a treatment for essential hypertension in 1964, and are still recommended as first choice because the cost for betablockers is low, which improves patient compliance. They act by binding to β_1 -receptors on the cardiac smooth muscle cells, which leads to decreased cardiac output. Most betablockers are not specific β_1 -receptor antagonists but binds to β_2 -receptors as well. The binding to
30 β_2 -receptors gives the opposite of the desired effect though inhibition of β_2 -receptors leads to vasoconstriction. This gives a side effect with cold hands and feet because most of the β_2 -receptors are located in the peripheral vascular arteries.

They have also been known to cause bronchospasms as well as some central nervous system side effects (nightmares, somnolence). They decrease insulin secretion, which makes them inappropriate to treat hypertensives with diabetes mellitus, and they can cause heart failure and peripheral artery obstructive disease (Velasco and
5 Rodriguez, Journal of Human Hypertension 10, Suppl. 1, S77-S80, 1996).

β -adrenoceptor agonists, such as dopamine and dobutamine are used to stimulate myocardial β_1 -adrenoceptors in the acute management of congestive heart failure. They act by increasing contractility and cardiac output.

A lot of different β_2 -agonists are used in the treatment of asthma. They
10 exert their primary effect on the β_2 -adrenergic receptor of bronchial smooth muscle, resulting in relaxation and bronchial dilatation. They also protect against bronchoconstrictor challenge (Hall, *supra*).

Thus, there is a clear need in the art for an improved understanding of the effects of betablockers on different subjects, and to predict which patients will have a
15 better response to treatment with betablockers.

Need for Effective Cardiovascular Status Assessment

The high morbidity and mortality associated with cardiovascular disease demonstrate a need in the art for methods and compositions that allow the determination
20 and/or prediction of the therapeutic regimen that will result in the most effective treatment outcome in a patient suffering from cardiovascular disease. This includes identification of individuals who are more or less responsive to particular therapeutic regimens, including, *e.g.*, particular drugs that are conventionally used to treat cardiovascular disease.

Furthermore, the heterogeneity in responses to cardiovascular therapies
25 emphasizes a need for another approach to rational drug development. In particular, populations that are identified as non-responsive to a particular therapeutic regimen can be identified for development and testing of alternative regimens. Thus, effective treatment regimens could be developed for a larger percentage of the affected population.

In summary, there is a need to reduce or eliminate trial and error in
30 selecting a therapeutic regimen for a particular individual. It would be desirable instead to predict whether a given individual will be responsive to, *e.g.*, a particular class of drugs or even to a particular drug or whether he/she is likely to suffer from adverse reactions or side-effects.

There is also a need in the art for methods and compositions that allow the identification of individuals having a predisposition to cardiovascular disease, such as, *e.g.*, myocardial infarction, hypertension, atherosclerosis, and stroke, to facilitate early intervention and disease prevention.

5 The present invention addresses these and other needs in the art by providing polymorphisms and polymorphic patterns that are characteristic of cardiovascular status, and by using these polymorphisms and patterns to prescribe or to develop more effective treatments or to assist in diagnosis.

Citation of any reference in this application should not be construed as an
10 admission that the reference is prior art to the invention. Each cited reference is incorporated herein by reference in its entirety.

SUMMARY OF THE INVENTION

In one aspect, the present invention provides reagents and methods for
15 predicting whether a particular therapeutic regime (such as a specific drug, a class of drugs or any other therapeutic regime, pharmacological or not) would be effective in improving a cardiovascular condition in a human individual, or would be ineffective for that purpose, or its use would be associated with adverse reactions or undesirable side-effects.

A particular advantage of the invention is that one or more polymorphic
20 markers provide a basis for predicting the outcome of a treatment regimen. By comparing a polymorphic pattern of a subject who requires treatment for a cardiovascular disorder, for example hypertension, with a reference pattern previously established to correlate with responsivity to the treatment regimen, a physician can predict whether a treatment plan, such as administration of an ACE inhibitor, is likely or not to be effective before
25 subjecting the subject to the treatment plan. For example, a comparison of the test polymorphic pattern from an individual with reference polymorphic patterns of individuals exhibiting differing responses to a particular therapeutic intervention can be used to predict the type or degree of responsivity of the individual to such intervention. The present invention thus represents a significant breakthrough in treating cardiovascular
30 pathologies in that it reduces or eliminates trial and error in selecting a treatment for a particular individual cardiovascular patient.

An additional advantage of the invention derives from the ability to eliminate subjects from clinical trials who are predictably non-responsive, or at risk for an

adverse response, to a particular treatment regimen. Furthermore, adverse results in an early trial can be evaluated to identify polymorphic patterns, so that the adverse results can be correlated with a sub-population of the test population permitting exclusion of such sub-population from the treatment group. The invention may thus ensure that a beneficial
5 drug can be approved for use in the appropriate population, and decrease the number of required patients and therefore the duration and cost of clinical trials. It may also lead to identification of another subgroup which can be the target for development of another therapeutic regimen.

All of the foregoing applications within the scope of the invention can be
10 deemed to be assessments of an individual's cardiovascular status, as the term is broadly defined below.

The foregoing methods of the invention are carried out by comparing a test polymorphic pattern established by at least one polymorphic position within a gene encoding ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor,
15 endothelin receptor, and β -adrenoceptor with a polymorphic pattern of a population of individuals exhibiting a predetermined responsivity to the regimen (reference pattern). If the test pattern matches the reference pattern, there is a statistically significant probability that the individual has the same cardiovascular status as that correlated with the reference pattern.

20 The polymorphic pattern preferably consists of at least two (and more preferably at least three) polymorphic positions, at least one of which is in the gene encoding a polypeptide from the group consisting of ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin receptor, and β -adrenoceptor, and a second polymorphism in a gene encoding a polypeptide selected from the group consisting
25 of ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin receptor, and β -adrenoceptor.

Additionally, the invention provides methods for assessing whether a particular individual has a genetic predisposition to a cardiovascular pathology. This aspect of the invention comprises comparing a test polymorphic pattern established by at
30 least one and preferably at least two and most preferably at least three polymorphic positions within a gene encoding ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin receptor, and β -adrenoceptor in conjunction with one or more, and preferably two or more, other polymorphic positions in ACE, AT1, AGT, renin,

aldosterone synthase, type-2 angiotensin II receptor, endothelin receptor, or β -adrenoceptor, with a polymorphic pattern of individuals exhibiting a predisposition to a cardiovascular syndrome. The conclusion drawn depends on whether the individual's polymorphism pattern matches the reference pattern.

- 5 The invention also provides an isolated nucleic acid having a sequence corresponding to part or all of the gene encoding ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin receptor, and β -adrenoceptor, the nucleic acid comprising a polymorphism in the ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin receptor, and β -adrenoceptor gene. In
10 preferred embodiments, the polymorphism, in combination with one or more other polymorphisms in the sequence of the same gene or a gene encoding a protein selected from the group consisting of human ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin receptor, and β -adrenoceptor, is predictive of a particular type or level of responsivity to a given treatment, or indicates a predisposition to
15 one or more clinical syndromes associated with cardiovascular disease, or both.

The isolated polymorphisms according to the invention (which are described using the numbering indicated in Table 1 below) include without limitation:

- Nucleic acids encoding renin having one or more polymorphic positions in the last exon of the gene (exon 10), which is a cytosine to thymine transition that creates a
20 premature stop codon at position 387.

Nucleic acids encoding aldosterone synthase promoter at position -344.

- Nucleic acids encoding β -adrenergic receptor-1 regulator region, at positions 2238, 2440, 2493, 2502, 2577, 2585, 2693, 2724, and 2757, as numbered in GenBank accession number X69168. In preferred embodiments, the bases at specific
25 positions are 2238 G, 2238 A, 2577 C, 2257 T, 2757 A, and 2757 G.

Nucleic acids encoding β -adrenergic receptor-1 coding region, at positions 231, 758, 1037, 1251, 1403, and 1528, as numbered in GenBank accession number J03019. In preferred embodiments, the bases at specific positions are 231 A, 231 G, 1251 C, 1251 G, 1403 A, 1403 G, 1528 C, and 1528 A.

- 30 Nucleic acids encoding β -adrenergic receptor-2 regulatory region, at positions 932, 934, 987, 1006, 1120, 1221, 1541, and 1568, as numbered in GenBank accession number M15169, or J02728, or M16106. In preferred embodiments, the bases

at specific positions are 934 A, 934 G, 987 C, 987 G, 1006 A, 1006 G, 1120 C, 1120 G, 1221 C, 1221 T, 1541 C, 1541 T, 1568 C, and 1568 T..

Nucleic acids encoding β -adrenergic receptor-2 coding region, at positions 839, 872, 1045, 1284, 1316, 1846, 1891, 2032, 2068, and 2070, as numbered in GenBank
5 accession number Y00106. In preferred embodiments, the bases at specific positions are 839 A, 839 G, 872 C, 872 G, 1045A, 1045 G, 1284 C, 1284 T, 1316 A, 1316 C, 1846 C, 1846 G, 2032 A, 2032 G, 2068 no insert, 2068 G, 2068 C, 2070 no insert, and 2070 C.

Nucleic acids encoding endothelin-A receptor at positions 969, 1005, 1146, and 2485, as numbered in GenBank accession number S57498. In preferred embodiments,
10 the bases at specific positions are 969 C, 969 T, 1005 A, 1005 G, 1146 A, 1146 G, 2485 T, and 2485 C.

Nucleic acids comprising polymorphisms present in other genes, which can be used in combination with a polymorphism from a gene encoding a polypeptide selected from the group consisting of ACE, AT1, AGT, renin, aldosterone synthase, type-2
15 angiotensin II receptor, endothelin receptor, and β -adrenoceptor to establish a polymorphism pattern, have been disclosed in International Patent Application No. PCT/IB98/00475, and include:

(i) Nucleic acids encoding ACE having one or more polymorphic positions at the position in the regulatory region numbered 5106; positions in the coding region
20 numbered 375, 582, 731, 1060, 2741, 3132, 3387, 3503, and 3906; and position 1451 all positions as numbered in GenBank entry X62855. In preferred embodiments, the sequences at the polymorphic positions in the ACE regulatory region are one or more of 5106C and 5106T; and the sequences at the polymorphic positions in the coding region are one or more of 375A, 375C, 582C, 582T, 731A, 731G, 1060G, 1060A, 2741G, 2741T,
25 3132C, 3132T, 3387T, 3387C, 3503G, 3503C, 3906G, and 3906A. The invention also encompasses a nucleic acid encoding a deletion of nucleotides 1451-1783 as numbered in GenBank entry X62855.

(ii) Nucleic acids encoding AGT having one or more polymorphic positions at positions in the regulatory region numbered 395, 412, 432, 449, 692, 839, 1007, 1072,
30 and 1204; positions in the coding region numbered 273, 912, 997, 1116, and 1174; and position 49 as numbered in GenBank entry M24688. In preferred embodiments, the sequences at the polymorphic positions in the AGT regulatory region are one or more of 395T, 395A, 412C, 412T, 432G, 432A, 449T, 449C, 692C, 692T, 839G, 839A, 1007G,

1007A, 1072G, 1072A, 1204C, and 1204A; the sequences at the polymorphic position in the coding region are one or more of 273C, 273T, 912C, 912T, 997G, 997C, 1116G, 1116A, 1174C and 1174A; and the sequence at position 49 in GenBank entry M24688 is either A or G.

- 5 (iii) Nucleic acids encoding AT1 having one or more polymorphic positions at positions in the regulatory region numbered 1427, 1756, 1853, 2046, 2354, 2355, and 2415; and the position in the coding region numbered 449. In preferred embodiments, the sequences at the polymorphic positions in the AT1 regulatory region are one or more of 1427A, 1427T, 1756T, 1756A, 1853T, 1853G, 2046T, 2046C, 2354A, 10 2354C, 2355G, 2355C, 2415A and 2415G; and the sequences at the polymorphic positions in the coding region are one or more of 449G, 449C, 678T, 678C, 1167A, 1167G, 1271A, and 1271C.

The invention also encompasses libraries of isolated nucleic acid sequences, such as arrays on a solid surface, wherein each sequence in the library 15 comprises a polymorphic position in the gene encoding ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin receptor, or β -adrenoceptor and other polymorphic positions in the other genes, including without limitation the polymorphic positions and sequences disclosed herein. Also provided are nucleic acid probes that hybridize specifically to the identified polymorphic positions; 20 peptides and polypeptides comprising polymorphic positions; and polymorphism-specific antibodies, *i.e.*, sequence-specific antibodies that bind differentially to polymorphic variants of the foregoing genes, that can be used to identify particular polymorphic variants.

In another aspect, the invention provides kits for the determination of 25 polymorphic patterns in an individual's genes. The kits comprise a means for detecting polymorphic sequences, including without limitation oligonucleotide probes that hybridize at or adjacent to the polymorphic positions and polymorphism-specific antibodies.

In yet another aspect, the invention provides nucleic acid and polypeptide targets for use in screening methods to identify candidate cardiovascular drugs. Nucleic acid targets may be, *e.g.*, DNA or RNA and are preferably at least about 10, and most 30 preferably at least about 15, residues in length and comprise one or more polymorphic positions in a gene encoding a polypeptide from the group consisting of ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin receptor, and β -

adrenoceptor. Peptide targets are at least about 5 amino acids in length and may be as large or larger than the full-length polypeptides.

5 **DETAILED DESCRIPTION OF THE INVENTION**

 The invention is based, in part, on the discovery that polymorphisms in certain genes in the RAAS, endothelin, and β -adrenoceptors pathways define polymorphism patterns that correlate with cardiovascular status. Most significantly, by comparing a test individual's polymorphism pattern with a reference polymorphism
10 pattern, which is a polymorphism pattern from a population of individuals with known cardiovascular status, one is able to predict whether the test individual has an increased likelihood to share the same cardiovascular status as that correlated with the reference polymorphism pattern. In particular, particular patterns correlate with responsiveness to ACE inhibitors, non-responsiveness to ACE inhibitors, and predisposition to
15 cardiovascular diseases or dysfunctions, including myocardial infarction and stroke.

 The invention provides a powerful predictive tool for clinical testing and treatment of cardiovascular disease. For clinical testing, the present invention permits smaller, more efficient clinical trials by identifying individuals who are likely to respond poorly to a treatment regimen and reducing the amount of uninterpretable data. By
20 evaluating a test individual's polymorphism pattern, a physician can prescribe a prophylactic or therapeutic regimen customized to that individual's cardiac status. Adverse responses to particular therapies can be avoided by excluding those individuals whose cardiovascular status puts them at risk for that therapy. Appropriate changes in lifestyle, including diet, environmental stress, and exercise levels can be prescribed for
25 individuals whose test polymorphic pattern matches a reference pattern that correlates with increased predisposition to cardiovascular disease.

Definitions

 "Cardiovascular status" as used herein refers to the physiological status of
30 an individual's cardiovascular system, as reflected in one or more status markers or indicators including genotype. Cardiovascular status shall be deemed to include without limitation not only the absence or presence of a pathology or disease in one or more components of the individual's cardiovascular system and the individual's predisposition

to developing such a condition, but also the individual's responsivity, *i.e.*, the ability or inability of the individual to respond (positively or negatively) to a particular prophylactic or therapeutic regimen or treatment for a cardiovascular condition, such as a drug or a class of drugs. A negative response includes one or more adverse reactions and side effects. Status markers include without limitation clinical measurements such as, *e.g.*, blood pressure, electrocardiographic profile, differentiated blood flow analysis, and the presence of increased levels of cellular proteins associated with a cardiovascular event. Examples of such proteins, also called diagnostic markers, which are important in cardiac events include myosin light chain, myosin heavy chain, myoglobin, troponin I, troponin T, CK-MB, etc. (see U.S. Patents No. 5,604,105 and No. 5,744,358). Status markers according to the invention are assessed using conventional methods well known in the art. Also included in the evaluation of cardiovascular status are quantitative or qualitative changes in status markers with time, such as would be used, *e.g.*, in the determination of an individual's response to a particular therapeutic regimen or of a predisposed individual's eventual development of a cardiovascular condition.

Examples of cardiovascular syndromes that are included in the foregoing definition of cardiovascular status include diagnosis of, or predisposition to, one or more cardiovascular syndromes, such as, *e.g.*, hypertension, acute myocardial infarction, silent myocardial infarction, unstable angina, stroke, and atherosclerosis. It will be understood that a diagnosis of a cardiovascular syndrome made by a medical practitioner encompasses not only clinical measurements but also medical judgment.

"Responsivity", as used herein, refers to the type and degree of response an individual exhibits to a particular therapeutic regimen, *i.e.*, the effect of a treatment on an individual. Responsivity breaks down into three major categories: therapeutic effect; no effect; and adverse effect. Naturally, there can be differing degrees of a therapeutic effect, *e.g.*, between full elimination and partial elimination of symptomology. In addition, adverse effects, or side effects, may be observed even though the treatment is beneficial, *i.e.*, therapeutically effective. Indeed, the present invention may permit identification of individuals with complex responsivity traits or patterns.

A "predisposition to develop a cardiovascular syndrome" refers to an increased likelihood, relative to the general population, to develop a cardiovascular syndrome, as defined above. A predisposition does not signify certainty, and development of the syndrome may be forestalled or prevented by prophylaxis, *e.g.*, adopting a modified

diet, exercise program, or treatment with gene therapy or pharmaceuticals. Naturally, an advantage of the present invention is that it permits identification of individuals who are, based on their genotype, predisposed to develop a cardiovascular syndrome, and for whom prophylactic intervention can be especially important.

5 A "polymorphism" as used herein denotes a variation in the nucleotide sequence of a gene in an individual. Genes that have different nucleotide sequences as a result of a polymorphism are "alleles". A "polymorphic position" is a predetermined nucleotide position within the sequence. In some cases, genetic polymorphisms are reflected by an amino acid sequence variation, and thus a polymorphic position can result
10 in location of a polymorphism in the amino acid sequence at a predetermined position in the sequence of a polypeptide. An individual "homozygous" for a particular polymorphism is one in which both copies of the gene contain the same sequence at the polymorphic position. An individual "heterozygous" for a particular polymorphism is one in which the two copies of the gene contain different sequences at the polymorphic
15 position.

 A "polymorphism pattern" as used herein denotes a set of one or more polymorphisms, including without limitation single nucleotide polymorphisms, which may be contained in the sequence of a single gene or a plurality of genes. In the simplest case, a polymorphism pattern can consist of a single nucleotide polymorphism in only one
20 position of one of two alleles of an individual. However, one has to look at both copies of a gene. A polymorphism pattern that is appropriate for assessing a particular aspect of cardiovascular status (*e.g.*, predisposition to hypertension) need not contain the same number (nor identity, of course) of polymorphisms as a polymorphism pattern that would be appropriate for assessing another aspect of cardiovascular status (*e.g.*, responsivity to
25 ACE inhibitors for control of hypertension). A "test polymorphism pattern" as used herein is a polymorphism pattern determined for a human subject of undefined cardiovascular status. A "reference polymorphism pattern" as used herein is determined from a statistically significant correlation of patterns in a population of individuals with pre-determined cardiovascular status.

30 "Nucleic acid" or "polynucleotide" as used herein refers to purine- and pyrimidine-containing polymers of any length, either polyribonucleotides or polydeoxyribonucleotides or mixed polyribo-polydeoxyribo nucleotides. Nucleic acids include without limitation single- and double-stranded molecules, *i.e.*, DNA-DNA, DNA-

RNA and RNA-RNA hybrids, as well as "protein nucleic acids" (PNA) formed by conjugating bases to an amino acid backbone. This also includes nucleic acids containing modified bases and non-naturally occurring phosphoester analog bonds, such as phosphorothioates and thioesters. The term nucleic acid molecule, and in particular DNA or RNA molecule, refers only to the primary and secondary structure of the molecule, and does not limit it to any particular tertiary forms. Thus, this term includes double-stranded DNA found, *inter alia*, in linear or circular DNA molecules (*e.g.*, restriction fragments), plasmids, and chromosomes. In discussing the structure of particular double-stranded DNA molecules, sequences may be described herein according to the normal convention of giving only the sequence in the 5' to 3' direction along the nontranscribed strand of DNA (*i.e.*, the strand having a sequence homologous to the mRNA). A "recombinant DNA molecule" is a DNA molecule that has undergone a molecular biological manipulation.

As used herein, the term "oligonucleotide" refers to a nucleic acid, generally of at least 10, preferably at least 15, and more preferably at least 20 nucleotides, that is hybridizable to a genomic DNA molecule, a cDNA molecule, or an mRNA molecule encoding a gene, cDNA, mRNA, or other nucleic acid of interest. Oligonucleotides can be labeled, *e.g.*, with ³²P-nucleotides or nucleotides to which a label, such as biotin, has been covalently conjugated. In one embodiment, a labeled oligonucleotide can be used as a probe to detect the presence of a nucleic acid. In another embodiment, oligonucleotides (one or both of which may be labeled) can be used as PCR primers, either for cloning full length or a fragment of a gene of interest, or to detect the presence of nucleic acids encoding the gene of interest. In a further embodiment, an oligonucleotide of the invention can form a triple helix with a double stranded sequence of interest in a DNA molecule. In still another embodiment, a library of oligonucleotides arranged on a solid support, such as a silicon wafer or chip, can be used to detect various polymorphisms of interest. Generally, oligonucleotides are prepared synthetically, preferably on a nucleic acid synthesizer. Accordingly, oligonucleotides can be prepared with non-naturally occurring phosphoester analog bonds, such as thioester bonds.

An "isolated" nucleic acid or polypeptide as used herein refers to a nucleic acid or polypeptide that is removed from its original environment (for example, its natural environment if it is naturally occurring). An isolated nucleic acid or polypeptide contains

less than about 50%, preferably less than about 75%, and most preferably less than about 90%, of the cellular components with which it was originally associated.

A nucleic acid or polypeptide sequence that is "derived from" a designated sequence refers to a sequence that corresponds to a region of the designated sequence. For
5 nucleic acid sequences, this encompasses sequences that are identical to or complementary to the sequence.

A "probe" refers to a nucleic acid or oligonucleotide that forms a hybrid structure with a sequence in a target nucleic acid due to complementarity of at least one sequence in the probe with a sequence in the target nucleic acid. Generally, a probe is
10 labeled so it can be detected after hybridization.

A nucleic acid molecule is "hybridizable" to another nucleic acid molecule, such as a cDNA, genomic DNA, or RNA, when a single stranded form of the nucleic acid molecule can anneal to the other nucleic acid molecule under the appropriate conditions of temperature and solution ionic strength (*see Sambrook et al.*, 1989, *Molecular Cloning: A*
15 *Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York). The conditions of temperature and ionic strength determine the "stringency" of the hybridization. For preliminary screening for homologous nucleic acids, low stringency hybridization conditions, corresponding to a T_m of 55°C, can be used, *e.g.*, 5x SSC, 0.1% SDS, 0.25% milk, and no formamide; or 30% formamide, 5x
20 SSC, 0.5% SDS). Moderate stringency hybridization conditions correspond to a higher T_m , *e.g.*, 40% formamide, with 5x or 6x SCC. High stringency hybridization conditions correspond to the highest T_m , *e.g.*, 50% formamide, 5x or 6x SCC. Hybridization requires that the two nucleic acids contain complementary sequences, although depending on the stringency of the hybridization, mismatches between bases are possible. The appropriate
25 stringency for hybridizing nucleic acids depends on the length of the nucleic acids and the degree of complementation, variables well known in the art. The greater the degree of similarity or homology between two nucleotide sequences, the greater the value of T_m for hybrids of nucleic acids having those sequences. The relative stability (corresponding to higher T_m) of nucleic acid hybridizations decreases in the following order: RNA:RNA,
30 DNA:RNA, DNA:DNA. For hybrids of greater than 100 nucleotides in length, equations for calculating T_m have been derived (*see Sambrook et al.*, *supra*, 9.50-9.51). For hybridization with shorter nucleic acids, *i.e.*, oligonucleotides, the position of mismatches becomes more important, and the length of the oligonucleotide determines its specificity

(see Sambrook *et al.*, *supra*, 11.7-11.8). A minimum length for a hybridizable nucleic acid is at least about 10 nucleotides; preferably at least about 15 nucleotides; and more preferably the length is at least about 20 nucleotides.

In a specific embodiment, the term "standard hybridization conditions" refers to a T_m of 55°C, and utilizes conditions as set forth above. In a preferred embodiment, the T_m is 60°C; in a more preferred embodiment, the T_m is 65°C. In a specific embodiment, "high stringency" refers to hybridization and/or washing conditions at 68°C in 0.2XSSC, at 42°C in 50% formamide, 4XSSC, or under conditions that afford levels of hybridization equivalent to those observed under either of these two conditions.

A "gene" for a particular protein as used herein refers to a contiguous nucleic acid sequence corresponding to a sequence present in a genome which comprises (i) a "coding (or transcribed) region," which comprises exons (*i.e.*, sequences encoding a polypeptide sequence, or "protein-coding" or "transcribed sequences"), introns, sequences at the junction between exons and introns, and 5' and 3' untranslated regions (uTRs); and (ii) regulatory sequences, which flank the coding region at both 5' and 3' termini. For example, the "ACE gene" as used herein encompasses the regulatory and coding regions of the human gene encoding angiotensin converting enzyme. Similarly, the "AGT gene" encompasses regulatory and coding regions of the human gene encoding angiotensinogen and the "AT1 gene" encompasses regulatory and coding regions of the human gene encoding type I angiotensin II receptor. Typically, regulatory sequences according to the invention are located 5' (*i.e.*, upstream) of the coding region segment. The reference sequences, obtained from GenBank, which were used in practicing the present invention are shown in Table 1.

Table 1. GenBank Accession Numbers

	Abbreviation	Compared master sequence	Numbering according to sequence entry in GenBank	SEQ ID NO.
5	AGT Regulatory Regions	X15323	X15323	122
	AGT Coding Region	M24686 (exon 2) M24687 (exon 3) M24688 (exon 4) M24689 (exon 5) X62855 (intron 16)	Protein-coding sequences from exon 2-5 were spliced together as described in the GenBank entries. Nucleotide 1 is assigned to the first nucleotide of the initiator methionine codon. X62855	123 124 125 126 127
	ACE Regulatory Region	X94359	X94359	128
10	ACE Coding Region	J04144	J04144 Nucleotide 1 is assigned to the first nucleotide of the initiator methionine codon.	129
	AT1 Regulatory Region	U07144	U07144	130
15	AT1 Coding Region	S80239 (exon 3) S77410 (exon 5)	The protein-coding sequence of S80239 was spliced to position 288 of entry S77410. Nucleotide 1 is assigned to the first nucleotide of the initiator methionine codon in entry S80239.	131 132
	Renin		M10030, X34914, X01391, X01734	133
	Aldosterone synthase		D13752	134
	AT2		U10273	135
20	β -adrenoceptor1 Regulatory Region (B1P)		X69168/g28421	136
25	β -adrenoceptor1 Coding Region (B1R)		J03019/g17899	137

β-adrenoceptor2 Regulatory Region (B2P)		M15169 J02728 M16106	138
β-adrenoceptor2 Coding Region (B2R)		Y00106/g29370	139
EndothelinA (ET _A) Receptor Coding		S57498	140

10

The present inventors have surprisingly and unexpectedly discovered the existence of genetic polymorphisms within the human gene encoding ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin receptor, or β-adrenoceptor which, singly or in combination, can be used to assess cardiovascular status, depending on which component of cardiovascular status is under evaluation. In accordance with the invention, the polymorphic pattern of the gene, alone or in combination with other genes encoding renin, ACE, AGT, AT1, AT2, aldosterone synthase, endothelin receptor, and β-adrenergic receptors 1 and 2 in an individual can predict the responsivity of the individual to particular therapeutic interventions and serve as an indicator of predisposition to various forms of cardiovascular disease. The invention provides methods for assessing cardiovascular status by detecting polymorphic patterns in an individual. The present invention also provides isolated nucleic acids derived from the gene which comprise these polymorphisms, including probes which hybridize specifically to polymorphic positions and primers that amplify the region of the gene in which the polymorphism is located; isolated polypeptides and peptides comprising polymorphic residues; and antibodies which specifically recognize ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin receptor, or β-adrenoceptor polypeptides containing one or more polymorphic amino acids.

30

Methods for Assessing Cardiovascular Status

The present invention provides diagnostic methods for assessing cardiovascular status in a human individual. The methods are carried out by comparing a polymorphic position or pattern ("test polymorphic pattern") within the individual's gene encoding ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin receptor, or β-adrenoceptor with the polymorphic patterns of humans

35

exhibiting a predetermined cardiovascular status ("reference polymorphic pattern"). If the cardiovascular status is the prediction of responsivity to a therapy, a single polymorphic position can provide a pattern for comparison. However, it is preferable to use more than one polymorphic position for the pattern to improve the accuracy of the prediction. If the

5 cardiovascular status is predisposition to a cardiovascular syndrome, at least two, and preferably at least three, polymorphic positions are used to make the pattern. In addition, other polymorphisms in genes encoding angiotensin converting enzyme (ACE), angiotensinogen (AGT), type 1 angiotensin II receptor (AT1), type 2 angiotensin II receptor, renin, aldosterone synthase, endothelin, receptor or β -adrenergic receptors 1 and

10 2 can be used to establish a polymorphic pattern for the individual.

For any meaningful prediction, the polymorphic pattern of the individual is identical to the polymorphic pattern of individuals who exhibit particular status markers, cardiovascular syndromes, and/or particular patterns of response to therapeutic interventions.

15 In one embodiment, the method involves comparing an individual's test polymorphic pattern with reference polymorphic patterns of individuals who have been shown to respond positively or negatively to a particular therapeutic regimen. Therapeutic regimen as used herein refers to treatments aimed at the elimination or amelioration of symptoms and events associated cardiovascular disease. Such treatments include without

20 limitation one or more of alteration in diet, lifestyle, and exercise regimen; invasive and noninvasive surgical techniques such as atherectomy, angioplasty, and coronary bypass surgery; and pharmaceutical interventions, such as administration of ACE inhibitors, angiotensin II receptor antagonists, diuretics, alpha-adrenoreceptor antagonists, cardiac glycosides, phosphodiesterase inhibitors, beta-adrenoreceptor antagonists, calcium channel

25 blockers, HMG-CoA reductase inhibitors, imidazoline receptor blockers, endothelin receptor blockers, and organic nitrites. Interventions with pharmaceutical agents not yet known whose activity correlates with particular polymorphic patterns associated with cardiovascular disease are also encompassed. The present inventors have discovered that particular polymorphic patterns correlate with an individual's responsivity to ACE

30 inhibitors (see, e.g., Example 3 below). It is contemplated, for example, that patients who are candidates for a particular therapeutic regimen will be screened for polymorphic patterns that correlate with responsivity to that particular regimen.

In another embodiment, the method involves comparing an individual's polymorphic pattern with polymorphic patterns of individuals who exhibit or have exhibited one or more markers of cardiovascular disease, such as, *e.g.*, high blood pressure, abnormal electrocardiographic profile, myocardial infarction, unstable angina, stroke, or atherosclerosis (see, *e.g.*, Example 2 below) and drawing analogous conclusions as to the individual's responsiveness to therapy, predisposition to developing a syndrome etc., as detailed above.

Identification of Polymorphic Patterns

10 In practicing the methods of the invention, an individual's polymorphic pattern can be established *e.g.* by obtaining DNA from the individual and determining the sequence at a predetermined polymorphic position or positions in a gene, or more than one gene.

The DNA may be obtained from any cell source. Non-limiting examples of cell sources available in clinical practice include without limitation blood cells, buccal cells, cervicovaginal cells, epithelial cells from urine, fetal cells, or any cells present in tissue obtained by biopsy. Cells may also be obtained from body fluids, including without limitation blood, saliva, sweat, urine, cerebrospinal fluid, feces, and tissue exudates at the site of infection or inflammation. DNA is extracted from the cell source or body fluid using any of the numerous methods that are standard in the art. It will be understood that the particular method used to extract DNA will depend on the nature of the source.

Determination of the sequence of the extracted DNA at polymorphic positions is achieved by any means known in the art, including but not limited to direct sequencing, hybridization with allele-specific oligonucleotides, allele-specific PCR, ligase-PCR, HOT cleavage, denaturing gradient gel electrophoresis (DGGE), and single-stranded conformational polymorphism (SSCP). Direct sequencing may be accomplished by any method, including without limitation chemical sequencing, using the Maxam-Gilbert method; by enzymatic sequencing, using the Sanger method; mass spectrometry sequencing; and sequencing using a chip-based technology. See, *e.g.*, Little *et al.*, Genet. Anal., 1996, 6:151. Preferably, DNA from a subject is first subjected to amplification by polymerase chain reaction (PCR) using specific amplification primers.

In an alternate embodiment, biopsy tissue is obtained from a subject. Antibodies that are capable of distinguishing between different polymorphic forms of

ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin receptor, or β -adrenoceptor are then applied to samples of the tissue to determine the presence or absence of a polymorphic form specified by the antibody. The antibodies may be polyclonal or monoclonal, preferably monoclonal. Measurement of specific antibody

5 binding to cells may be accomplished by any known method, *e.g.*, quantitative flow cytometry, or enzyme-linked or fluorescence-linked immunoassay. The presence or absence of a particular polymorphism or polymorphic pattern, and its allelic distribution (*i.e.*, homozygosity vs. heterozygosity) is determined by comparing the values obtained from a patient with norms established from populations of patients having known

10 polymorphic patterns.

In another alternate embodiment, RNA is isolated from biopsy tissue using standard methods well known to those of ordinary skill in the art such as guanidium thiocyanate-phenol-chloroform extraction (Chomocyznski *et al.*, *Anal. Biochem.*, 1987, 162:156). The isolated RNA is then subjected to coupled reverse transcription and

15 amplification by polymerase chain reaction (RT-PCR), using specific oligonucleotide primers that are specific for a selected polymorphism. Conditions for primer annealing are chosen to ensure specific reverse transcription and amplification; thus, the appearance of an amplification product is diagnostic of the presence of a particular polymorphism. In another embodiment, RNA is reverse-transcribed and amplified, after which the amplified

20 sequences are identified by, *e.g.*, direct sequencing. In still another embodiment, cDNA obtained from the RNA can be cloned and sequenced to identify a polymorphism.

Establishing Reference Polymorphism Patterns

In practicing the present invention, the distribution of polymorphic patterns

25 in a large number of individuals exhibiting particular cardiovascular status is determined by any of the methods described above, and compared with the distribution of polymorphic patterns in patients that have been matched for age, ethnic origin, and/or any other statistically or medically relevant parameters, who exhibit quantitatively or qualitatively different cardiovascular status. Correlations are achieved using any method

30 known in the art, including nominal logistic regression or standard least squares regression analysis. In this manner, it is possible to establish statistically significant correlations between particular polymorphic patterns and particular cardiovascular statuses. It is further possible to establish statistically significant correlations between particular

polymorphic patterns and changes in cardiovascular status such as, would result, *e.g.*, from particular treatment regimens. Thus, it is possible to correlate polymorphic patterns with responsivity to particular treatments.

5 A statistically significant correlation preferably has a "p" value of less than or equal to 0.05. Any standard statistical method can be used to calculate these values, such as the normal Student's T Test, or Fischer's Exact Test.

The identity and number of polymorphisms to be included in a reference pattern depends not only on the prevalence of a polymorphism and its predictive value for the particular use, but also on the value of the use and its requirement for accuracy of
10 prediction. The greater the predictive value of a polymorphism, the lower the need for inclusion of more than one polymorphism in the reference pattern. However, if a polymorphism is very rare, then its absence from an individual's pattern might provide no indication as to whether the individual has a particular status. Under these circumstances, it might be advisable to select instead two or more polymorphisms which are more
15 prevalent. Even if none of them has a high predictive value on its own, the presence of both (or all three) of them might be sufficiently predictive for the particular purpose.

If for example the use for a reference pattern is prediction of response to a drug, and among the afflicted population only a 30% response to the drug is observed, the reference pattern need only permit selection of a population that improves the response
20 rate by 10% to provide a significant improvement in the state of the art. On the other hand, if the use for the reference pattern is selection of subjects for a particular clinical study, the pattern should be as selective as possible and should therefore include a plurality of polymorphisms that together provide a high predictive accuracy for the intended response.

25 In establishing reference polymorphism patterns, it is desirable to use a defined population. For example, tissue libraries collected and maintained by state or national departments of health can provide a valuable resource, since genotypes determined from these samples can be matched with medical history, and particularly cardiovascular status, of the individual. Such tissue libraries are found, for example, in
30 Sweden, Iceland, Norway, and Finland. As can be readily understood by one of ordinary skill in the art, specific polymorphisms may be associated with a closely linked population. However, other polymorphisms in the same gene may correlate with cardiovascular status of other genetically related populations. Thus, in addition to the

specific polymorphisms provided in the instant application, the invention identifies genes in which any polymorphisms can be used to establish reference and test polymorphism patterns for evaluating cardiovascular status of individuals in the population.

In a specific embodiment, DNA samples can be obtained from a well defined population, such as 277 Caucasian males born in Uppsala, Sweden between 1920 and 1924. In a specific embodiment, such individuals are selected for the test population based on their medical history, *i.e.*, they were either (i) healthy, with no signs of cardiovascular disease (100); or (ii) had suffered one of acute myocardial infarction (68), silent myocardial infarction (34), stroke (18), stroke and acute myocardial infarction (19), or high blood pressure at age 50 (39). DNA samples are obtained from each individual.

In a specific embodiment, DNA sequence analysis can be carried out by: (i) amplifying short fragments of each of the genes using polymerase chain reaction (PCR) and (ii) sequencing the amplified fragments. The sequences obtained from each individual can then be compared with the first known sequences, *e.g.*, as set forth in Table 1, to identify polymorphic positions.

Comparing Test Patterns to Reference Patterns

As noted above, the test pattern from an individual can be compared to a reference pattern established for a predetermined cardiovascular status. Identity between the test pattern and the reference pattern means that the tested individual has a probability of having the same cardiovascular status as that represented by the reference pattern. As discussed above, this probability depends on the prevalence of the polymorphism and the statistical significance of its correlation with a cardiovascular status.

25

Polymorphic Positions

Polymorphic positions in the genes encoding ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin receptor, or β -adrenoceptor which are encompassed by the invention can be identified by determining the DNA sequence of all or part of the gene in a multiplicity of individuals in a population. DNA sequence determination may be achieved using any conventional method, including, *e.g.*, chemical or enzymatic sequencing.

The polymorphic positions of genes for use in the invention include without limitation those listed below, whose numbering corresponds to the GenBank sequences listed in Table 1.

- (i) ACE: positions in the regulatory region (designated ACR) numbered 5106, 5349, and 5496; positions in the coding region (designated ACE) numbered 375, 582, 731, 1060, 1215, 2193, 2328, 2741, 3132, 3387, 3503, and 3906; and position 1451 as numbered in GenBank entry X62855.
- (ii) AGT: positions in the regulatory region (designated AGR) numbered 395, 412, 432, 449, 692, 839, 1007, 1072, 1204, and 1218; positions in the coding region (designated AGT) numbered 273, 620, 803, 912, 997, 1116, and 1174; and position 49 as numbered in GenBank entry M24688.
- (iii) AT1: positions in the regulatory region (designated ATR) numbered 1427, 1756, 1853, 2046, 2354, 2355, and 2415; and positions in the coding region (designated AT1) numbered 449, 678, 1167, and 1271.
- (iv) Renin: A mutant renin gene in familial elevation of prorenin, a point mutation in the last exon of the gene (exon 10), has been identified (Villard *et al.*, J. Biol. Chem., 1994, 269:30307-12). A cytosine to thymine transition creates a premature stop codon at position 387 resulting in a truncated form of renin with 20 amino acids deleted from the carboxyl terminus.
- (v) Aldosterone synthase: A position in the promoter region of aldosterone synthase, position -344 (with the initiation codon starting at 1) has been reported by Cambien *et al.* at the International Meeting on Hypertension held in Amsterdam in June 1998.
- β -adrenergic receptor-1, positions in the regulatory region (designated BP1) numbered 2238, 2440, 2493, 2502, 2577, 2585, 2693, 2724, and 2757; and positions in the coding region (designated BR1) numbered 231, 758, 1037, 1251, 1403, and 1528.
- β -adrenergic receptor-2, positions in the regulatory region (designated B2P) numbered 932, 934, 987, 1006, 1120, 1221, 1541, and 1568; and positions in the coding region (designated B2R) numbered 839, 872, 1045, 1284, 1316, 1846, 1891, 2032, 2068, and 2070.
- Endothelin receptor type A coding region (designated ET_A) numbered 969, 1005, 1146, and 2485.

In preferred embodiments, the base at each of the above polymorphic positions is one of:

- (i) ACE Regulatory Region: 5106C, 5106T, 5349A, 5349T, 5496T, and 5496C;
- 5 (ii) ACE Coding Region: 375A, 375C, 582C, 582T, 731A, 731G, 1060G, 1060A, 1215C, 1215T, 2193G, 2193A, 2328A, 2328G, 2741G, 2741T, 3132C, 3132T, 3387T, 3387C, 3503G, 3503C, 3906G, and 3906A; and a deletion of nucleotides 1451-1783 as numbered in GenBank entry X62855;
- (iii) AGT Regulatory Region: 395T, 395A, 412C, 412T, 432G, 432A, 10 449T, 449C, 692C, 692T, 839G, 839A, 1007G, 1007A, 1072G, 1072A, 1204C, 1204A, 1218A, 1218G;
- (iv) AGT Coding Region: 273C, 273T, 620C, 620T, 803T, 803C, 912C, 912T, 997G, 997C, 1116G, 1116A, 1174C, and 1174A; and A or G at position 49 in GenBank entry M24688;
- 15 (v) AT1 Regulatory Region: 1427A, 1427T, 1756T, 1756A, 1853T, 1853G, 2046T, 2046C, 2354A, 2354C, 2355G, 2355C, 2415A and 2415G; and
- (vi) AT1 Coding Region: 449G, 449C, 678T, 678C, 1167A, 1167G, 1271A, and 1271C.
- (vii) β -adrenergic receptor-1 regulatory region: 2238 G, 2238 A, 2577 C, 20 2257 T, 2757 A, and 2757 G.
- (viii) β -adrenergic receptor-1 coding region: 231 A, 231 G, 758 C, 758 T, 1251 C, 1251 G, 1403 A, 1403 G, 1528 C, and 1528 A.
- (ix) β -adrenergic receptor-2 regulatory region: 934 A, 934 G, 987 C, 987 G, 1006 A, 1006 G, 1120 C, 1120 G, 1221 C, 1221 T, 1541 C, 1541 T, 1568 C, and 25 1568 T.
- (x) β -adrenergic receptor-2 coding region: 839 A, 839 G, 872 C, 872 G, 1045A, 1045 G, 1284 C, 1284 T, 1316 A, 1316 C, 1846 C, 1846 G, 2032 A, 2032 G, 2068 no insert, 2068 G, 2068 C, 2070 no insert, and 2070 C.
- (xi) Endothelin receptor type A: 969 C, 969 T, 1005 A, 1005 G, 1146 30 A, 1146 G, 2485 T, and 2485 C.

An individual may be homozygous or heterozygous for a particular polymorphic position (see, e.g., Table 6 below).

Non-limiting examples of polymorphic patterns comprising one or more polymorphism in ACE, AGT, and/or AT1 genes according to the invention include the following, which were correlated with an increased incidence of clinical signs of cardiovascular disease:

- 5 ACR 5349 A/T, AGR 1218 A; ACR 5496 C, AGR 1204 A/C; ACR 5496 C/T, AGR 1218 A, AGT 620 C/T; ACE 2193 A, AGR 1204 C, ACE 2328 G; ACE 2193 A, AGR 1204 A/C; ACE 3387 T, AGR 1218 A; ACE 3387 T, AGT 620 C/T; AGR 1204 A/C, AT1 678 C/T; AGR 1204 A/C, AT1 1271 A/C; ACE 1215 C, AGR 1204 A/C; AGR 1204 A/C, AT1 1167 A, ACE 3906 A/G; AGR 1204 A, AGT 620 C, AT1 1271 A, AT1
10 1167 A, AGR 395 A/T; AGR 1204 A/C, AGT 620 C/T, AT1 1271 A/C, AT1 1167 A, AGR 395 T; AGR 1204 A/C, AGT 620 C/T, AT1 1271 A/C, AT1 1167 A/G, AGR 395 T; AGR 1204 A, AT1 678 C, AT1 1167 A, AGR 395 A/T; AGR 1204 A/C, AT1 678 C/T, AT1 1167 A, AGR 395 T; AGT 620 C/T, AT1 1271 A/C, AT1 1167 A, AGR 395 T; AGT 620 C/T, AT1 1271 A/C, AT1 1167 A/G, AGR 395 T; AGT 620 C, AT1 1271 A, AT1
15 1167 A, AGR 395 A/T; AGT 620 C, AT1 678 A, AT1 1167 A, AGR 395 A/T; AGT 620 C/T, AT1 678 C/T; AT1 1167 A, AGR 395 T; ACE 2193 A, AGR 1218 A, AGT 803 A; ACE 2193 A, AGT 620 C/T; ACE 2328 G, AGT 620 C/T; ACE 3387 T, AGR 1204 A/C; ACE 2193 A, ACE 2328 G, AGR 1204 C; ACE 2193 A/G, AGR 1072 G/G, AT1 1167 A/A. Additional polymorphism patterns are shown in the Tables in Examples 4 and 5,
20 below.

Polymorphism Patterns Correlated With ACE Inhibitor Responsiveness

The following table lists a set of polymorphism patterns that have been found to correlate with responsiveness to ACE inhibitor treatment:

25 Response to ACE-Inhibitor Treatment

30	Position 1	Genotype 1	Position 2	Genotype 2	Position 3	Genotype 3
	ACE:2193:	A/G	AGR:1072:	G/G	AT1:1167:	A/A
	ACE:2193:	A/G	AGR:1072:	G/G	ACE:1060:	G/G
	ATR:2354:	A/A	AT1:678:	C/T	AT1:1167:	A/A
	ACR:5496:	C/T	AGR:1204:	A/A	AGR:839:	G/G
	ACR:5496:	C/T	AGR:1204:	A/A	AGT:620:	C/C
	ACR:5496:	C/T	ACE:1060:	G/G	AGR:449:	C/T

	Position 1	Genotype 1	Position 2	Genotype 2	Position 3	Genotype 3
	ACR:5496:	C/T	ACE:1215:	C/T	AGR:1204:	A/A
	ACR:5496:	C/T	ACE:3906:	A/G	AGR:1218:	A/G
	ACR:5496:	C/T	AGR:449:	C/T	AGR:839:	G/G
	ACR:5496:	C/T	AGR:449:	C/T	AT1:1271:	A/A
5	ACR:5496:	C/T	AGR:449:	C/T	AGR:1072:	G/G
	ACR:5496:	C/T	AGR:1072:	G/G	AGR:1204:	A/A
	ACE:1060:	G/G	AGR:449:	C/T	AGT:620:	C/C
	ACE:1060:	G/G	AGR:1007:	G/G	AT1:678:	C/T
	ACE:1215:	C/T	ACE:3906:	A/G	AGR:1218:	A/G
10	AGR:449:	C/T	AGR:1204:	A/A	AGT:620:	C/C
	AGR:449:	C/T	AGR:1204:	A/A	AGT:1116:	G/G
	AGR:449:	C/T	AGR:839:	G/G	AGT:620:	C/C
	AGR:449:	C/T	AGT:620:	C/C	AGT:1116:	G/G
	AGR:449:	C/T	AGT:1116:	G/G	AT1:1271:	A/A
15	AGR:449:	C/T	AGR:1072:	G/G	AGT:620:	C/C
	AGR:1007:	G/C	AGR:1072:	G/G	AT1:678:	C/T
	AGR:1072:	G/G	AGR:1204:	A/A	AGT:1116:	G/G
	ACR:5496:	C/T	AGR:1204:	A/A	AT1:1167:	A/A
20	AGR:1007:	G/G	AGR:1204:	A/A	AGT:1116:	G/G

Polymorphism Patterns of Correlated With ACE Inhibitor Non-responsiveness

The following table lists a set of polymorphism patterns that have been found to correlate with non-responsiveness to ACE inhibitor treatment:

Non-Response to ACE-Inhibitor Treatment

25	Position 1	Genotype 1	Position 2	Genotype 2	Position 3	Genotype 3
	ACE:1060:	<u>G/G</u>	AGR:1204:	<u>A/C</u>	AGT:620:	<u>C/C</u>
	ACE:1215:	<u>C/T</u>	ACE:2193:	<u>A/G</u>	AGR:1204:	<u>A/C</u>
	ACE:2193:	<u>A/G</u>	ATR:2046:	<u>C/T</u>	AT1:678:	<u>C/C</u>

Position 1	Genotype 1	Position 2	Genotype 2	Position 3	Genotype 3
ACE:2193:	A/G	AT1:678:	C/C	AT1:1271:	A/C

Polymorphism Patterns Correlated With Predisposition to MI

- The following table lists a set of polymorphism patterns that have been found to correlate with predisposition to myocardial infarction:

Predisposition to MI

Position 1	Genotype 1	Position 2	Genotype 2	Position 3	Genotype 3
ACE:2193:	A/A	AGR:1204:	A/C	-	-
AGR:449:	T/T	AGR:1204:	A/C	AT1:1271:	A/C
ACE:2193:	A/G	AGR:620:	C/T	AGR:1116	G/G
ACE:2193:	A/G	AGR:449:	T/T	AT1:1271:	A/C
AGR:5349:	A/A	AGR:449:	T/T	AT1:1271:	A/C
ACE:2193:	A/A	AGR:620:	C/T	AT1:1271:	A/A

Polymorphism Patterns Correlated With Predisposition to Stroke

- The following table lists a set of polymorphism patterns that have been found to correlate with predisposition to stroke:

Predisposition to Stroke

Position 1	Genotype 1	Position 2	Genotype 2	Position 3	Genotype 3
ACE:2193:	A/A	AGR:395:	A/T	-	-
AGR:1007:	A/G	AGR:1072:	G/G	AT1:1167:	A/A
AGR:395:	A/T	AGR:1072:	G/G	AGR:1218:	A/G

Polymorphism Patterns in β -Adrenergic Receptor Genes

- The following table lists a set of polymorphism patterns that have been found in β -adrenergic receptor genes:

Positions carrying genetic variation in the Beta adrenergic receptors 1 and

2.

	B1P	B1R	B2R	B2P
	2238 C/A	231 A/G	839 A/G	932
	2440	758 C/T 1037	872 C/G	934 A/G
	2493	1251 C/G	1045 A/G	987 C/G 1006 A/G
5	2502	1403 A/G	1284 C/T	1120 C/G
	2577 C/T	1528 C/A	1316 A/C	1221 C/T
	2585		1846 C/G	1541 C/T
	2693		1891	1568 C/T
	2724		2032 A/G	
10	2757 A/G		2068 no insert, G/G, C/G 2070 no insert /C	

B1P: Beta adrenergic receptor 1, regulatory promoter region.

B1R: Beta adrenergic receptor 1, coding region.

B2P: Beta adrenergic receptor 2, regulatory promoter region.

15 B2R: Beta adrenergic receptor 2, coding region.

A number of different polymorphisms have been identified in the type 2 β -adrenoceptor. All of these differed from the wild type sequence by a single base change. Four of the polymorphisms alter the amino acid sequence of the receptor protein (Hall, Thorax, 1996, 51:351-353). The amino acid sequence modifications are described in
20 greater detail below:

Arg16-Gly: The Gly16 variant undergoes an enhanced agonist-promoted down regulation as compared to wild type but the coupling to adenylyl cyclase and agonist binding are maintained (Liggett, Am. J. Respir. Crit. Care Med., 1997, 156:S 156-S162).

Gln27-Glu: The Glu27 variant displays very little agonist-promoted
25 downregulation and the coupling to adenylyl cyclase and agonist binding are maintained (*id.*).

Val34-Met: Met34 is very rare. No altering of receptor function has been found (*id.*).

Thr164-Ile: Uncommon (about 5%). The Ile164 variant shows depressed coupling to adenylyl cyclase and decreased affinities for agonists with hydroxyl groups on their β -carbons, such as epinephrine, norepinephrine, and isoproterenol compared to wild type (*id.*).

5 The polymorphism at nucleic acid 523 (CGG-AGG) might be linked with one of the other functional polymorphisms (*id.*).

There are no differences in frequency of these polymorphisms between the normal group and those with asthma but they have been correlated to differences in response to treatment with agonists in asthma, *e.g.*, the Gly16 variant undergoes an
 10 enhanced agonist-promoted downregulation compared to wild type (*id.*).

Polymorphism Patterns in Endothelin Receptor Type A Gene

The following table lists a set of polymorphism patterns that have been found in the coding region of the endothelin receptor type A gene:

15	Position 969	Position 1005	Position 1146	Position 2485
	C/C	A/A	A/A	T/T
	C/T	A/G	A/G	T/C
	T/T	G/G		

Isolated Polymorphic Nucleic Acids, Vectors, Probes & Primers and Arrays

Vectors for Expression of Polymorphic Variants

The present invention provides isolated nucleic acids comprising the polymorphic positions described above for the human genes encoding ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin receptor, and β -
 25 adrenoceptor; vectors comprising the nucleic acids; and transformed host cells comprising the vectors. The invention also provides probes which are useful for detecting these polymorphisms.

The nucleic acids encoding a gene comprising a polymorphism that is useful for determining cardiovascular status of an individual is particularly valuable for
 30 screening, whether by direct screening of the nucleic acid with the polymorphism, or by screening the polypeptide expressed by that nucleic acid.

In practicing the present invention, many conventional techniques in molecular biology, microbiology, and recombinant DNA, are used. Such techniques are well known and are explained fully in, for example, Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Second Edition, 1989 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York); *DNA Cloning: A Practical Approach*, Volumes I and II, 1985 (D.N. Glover ed.); *Oligonucleotide Synthesis*, 1984, (M.L. Gait ed.); *Nucleic Acid Hybridization*, 1985, (Hames and Higgins); Ausubel *et al.*, *Current Protocols in Molecular Biology*, 1997 (John Wiley and Sons); and *Methods in Enzymology* Vol. 154 and Vol. 155 (Wu and Grossman, and Wu, eds., respectively).

10 Insertion of nucleic acids (typically DNAs) comprising the sequences of the present invention into a vector is easily accomplished when the termini of both the DNAs and the vector comprise compatible restriction sites. If this cannot be done, it may be necessary to modify the termini of the DNAs and/or vector by digesting back single-stranded DNA overhangs generated by restriction endonuclease cleavage to produce blunt
15 ends, or to achieve the same result by filling in the single-stranded termini with an appropriate DNA polymerase.

 Alternatively, any site desired may be produced, *e.g.*, by ligating nucleotide sequences (linkers) onto the termini. Such linkers may comprise specific oligonucleotide sequences that define desired restriction sites. Restriction sites can also be generated by
20 the use of the polymerase chain reaction (PCR). *See, e.g.*, Saiki *et al.*, *Science*, 1988, 239:48. The cleaved vector and the DNA fragments may also be modified if required by homopolymeric tailing.

 The nucleic acids may be isolated directly from cells or may be chemically synthesized using known methods. Alternatively, the polymerase chain reaction (PCR)
25 method can be used to produce the nucleic acids of the invention, using either chemically synthesized strands or genomic material as templates. Primers used for PCR can be synthesized using the sequence information provided herein and can further be designed to introduce appropriate new restriction sites, if desirable, to facilitate incorporation into a given vector for recombinant expression.

30 The nucleic acids of the present invention may be flanked by native gene sequences, or may be associated with heterologous sequences, including promoters, enhancers, response elements, signal sequences, polyadenylation sequences, introns, 5'- and 3'- noncoding regions, and the like.

The invention also provides nucleic acid vectors comprising the disclosed genes or derivatives or fragments thereof. A large number of vectors, including plasmid and fungal vectors, have been described for replication and/or expression in a variety of eukaryotic and prokaryotic hosts, and may be used for gene therapy as well as for simple cloning or protein expression. Non-limiting examples of suitable vectors include without limitation pUC plasmids, pET plasmids (Novagen, Inc., Madison, WI), or pRSET or pREP (Invitrogen, San Diego, CA), and many appropriate host cells, using methods disclosed or cited herein or otherwise known to those skilled in the relevant art. The particular choice of vector/host is not critical to the practice of the invention.

Suitable host cells may be transformed/transfected/infected as appropriate by any suitable method including electroporation, CaCl_2 mediated DNA uptake, calcium phosphate precipitation, fungal or viral infection, lipofection, microinjection, microprojectile, or other established methods. Appropriate host cells included bacteria, archebacteria, fungi, especially yeast, and plant and animal cells, especially mammalian cells. A large number of transcription initiation and termination regulatory regions have been isolated and shown to be effective in the transcription and translation of heterologous proteins in the various hosts. Examples of these regions, methods of isolation, manner of manipulation, *etc.* are known in the art. Under appropriate expression conditions, host cells can be used as a source of recombinantly produced ACE-, AGT-, or AT1-derived peptides and polypeptides.

Nucleic acids encoding ACE-, AGT-, or AT1-derived gene sequences may also be introduced into cells by recombination events. For example, such a sequence can be introduced into a cell and thereby effect homologous recombination at the site of an endogenous gene or a sequence with substantial identity to the gene. Other recombination-based methods such as nonhomologous recombinations or deletion of endogenous genes by homologous recombination may also be used.

Oligonucleotides

The nucleic acids of the present invention find use as probes for the detection of genetic polymorphisms, as primers for the expression of polymorphisms, or in molecular library arrays for high throughput screening.

Probes in accordance with the present invention comprise without limitation isolated nucleic acids of about 10 - 100 bp, preferably 15-75 bp and most

preferably 17-25 bp in length, which hybridize at high stringency to one or more of the gene-derived polymorphic sequences disclosed herein or to a sequence immediately adjacent to a polymorphic position. Furthermore, in some embodiments a full-length gene sequence may be used as a probe. In one series of embodiments, the probes span the
5 polymorphic positions in the genes disclosed above. In another series of embodiments, the probes correspond to sequences immediately adjacent to the polymorphic positions.

The oligonucleotide nucleic acids may also be modified by many means known in the art. Non-limiting examples of such modifications include methylation, "caps", substitution of one or more of the naturally occurring nucleotides with an analog,
10 internucleotide modifications such as, for example, those with uncharged linkages (*e.g.*, methyl phosphonates, phosphotriesters, phosphoroamidates, carbamates, *etc.*) and with charged linkages (*e.g.*, phosphorothioates, phosphorodithioates, *etc.*). Nucleic acids may contain one or more additional covalently linked moieties, such as, for example, proteins (*e.g.*, nucleases, toxins, antibodies, signal peptides, poly-L-lysine, *etc.*), intercalators (*e.g.*,
15 acridine, psoralen, *etc.*), chelators (*e.g.*, metals, radioactive metals, iron, oxidative metals, *etc.*), and alkylators. PNAs are also included. The nucleic acid may be derivatized by formation of a methyl or ethyl phosphotriester or an alkyl phosphoramidate linkage. Furthermore, the nucleic acid sequences of the present invention may also be modified with a label capable of providing a detectable signal, either directly or indirectly.
20 Exemplary labels include radioisotopes, fluorescent molecules, biotin, and the like.

PCR amplification of gene segments that contain a polymorphism provides a powerful tool for detecting the polymorphism. The oligonucleotides of the invention can also be used as PCR primers to amplify gene segments containing a polymorphism of interest. The amplified segment can be evaluated for the presence or absence of a
25 polymorphism by restriction endonuclease activity, SSCP, or by direct sequencing. In another embodiment, the primer is specific for a polymorphic sequence on the gene. If the polymorphism is present, the primer can hybridize and DNA will be produced by PCR. However, if the polymorphism is absent, the primer will not hybridize, and no DNA will be produced. Thus, PCR can be used to directly evaluate whether a polymorphism is
30 present or absent.

Molecular library arrays of oligonucleotides (including oligonucleotides with modifications as described above) are another powerful tool for rapidly assessing whether one or more polymorphisms are present in a *ACE*, *ATI*, *AGT*, *renin*, *aldosterone*

synthase, type-2 angiotensin II receptor, endothelin receptor, or β -adrenoceptor gene, preferably in combination with other genes. Molecular library arrays are disclosed in US Patents No. 5,677,195, No. 5,599,695, No. 5,545,531, and No. 5,510,270.

5 **Polymorphic Polypeptides and Polymorphism-Specific Antibodies**

 The present invention encompasses isolated peptides and polypeptides encoded by all or a portion of a gene encoding a polypeptide selected from the group consisting of ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin receptor, and β -adrenoceptor, comprising polymorphic positions disclosed
10 above. In one preferred embodiment, the peptides and polypeptides are useful screening targets to identify cardiovascular drugs. In another preferred embodiment, the peptides and polypeptides are capable of eliciting antibodies in a suitable host animal that react specifically with a polypeptide comprising the polymorphic position and distinguish it from other polypeptides having a different amino acid sequence at that position.

15 Polypeptides according to the invention are preferably at least five or more residues in length, preferably at least fifteen residues. Methods for obtaining these polypeptides are described below. Many conventional techniques in protein biochemistry and immunology are used. Such techniques are well known and are explained in *Immunochemical Methods in Cell and Molecular Biology*, 1987 (Mayer and Waler, eds;
20 Academic Press, London); Scopes, *Protein Purification: Principles and Practice*, Second Edition 1987 (Springer-Verlag, N.Y.) and *Handbook of Experimental Immunology*, Volumes I-IV 1986 (Weir and Blackwell eds.).

 Nucleic acids comprising protein-coding sequences can be used to direct the recombinant expression of ACE, AT1, AGT, renin, aldosterone synthase, type-2
25 angiotensin II receptor, endothelin receptor, or β -adrenoceptor-derived polypeptides in intact cells or in cell-free translation systems. The known genetic code, tailored if desired for more efficient expression in a given host organism, can be used to synthesize oligonucleotides encoding the desired amino acid sequences. The polypeptides may be isolated from human cells, or from heterologous organisms or cells (including, but not
30 limited to, bacteria, fungi, insect, plant, and mammalian cells) into which an appropriate protein-coding sequence has been introduced and expressed. Furthermore, the polypeptides may be part of recombinant fusion proteins.

Peptides and polypeptides may be chemically synthesized by commercially available automated procedures, including, without limitation, exclusive solid phase synthesis, partial solid phase methods, fragment condensation or classical solution synthesis. The polypeptides are preferably prepared by solid phase peptide synthesis as
5 described by Merrifield, J. Am. Chem. Soc., 1963, 85:2149.

Methods for polypeptide purification are well-known in the art, including, without limitation, preparative disc-gel electrophoresis, isoelectric focusing, HPLC, reversed-phase HPLC, gel filtration, ion exchange and partition chromatography, and countercurrent distribution. For some purposes, it is preferable to produce the polypeptide
10 in a recombinant system in which the protein contains an additional sequence tag that facilitates purification, such as, but not limited to, a polyhistidine sequence. The polypeptide can then be purified from a crude lysate of the host cell by chromatography on an appropriate solid-phase matrix. Alternatively, antibodies produced against ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin receptor, or
15 β -adrenoceptor or against peptides derived therefrom, can be used as purification reagents. Other purification methods are possible.

The present invention also encompasses derivatives and homologues of the polypeptides. For some purposes, nucleic acid sequences encoding the peptides may be altered by substitutions, additions, or deletions that provide for functionally equivalent
20 molecules, *i.e.*, function-conservative variants. For example, one or more amino acid residues within the sequence can be substituted by another amino acid of similar properties, such as, for example, positively charged amino acids (arginine, lysine, and histidine); negatively charged amino acids (aspartate and glutamate); polar neutral amino acids; and non-polar amino acids.

25 The isolated polypeptides may be modified by, for example, phosphorylation, sulfation, acylation, or other protein modifications. They may also be modified with a label capable of providing a detectable signal, either directly or indirectly, including, but not limited to, radioisotopes and fluorescent compounds.

The present invention also encompasses antibodies that specifically
30 recognize the polymorphic positions of the invention and distinguish a peptide or polypeptide containing a particular polymorphism from one that contains a different sequence at that position. Such polymorphic position-specific antibodies according to the present invention include polyclonal and monoclonal antibodies. The antibodies may be

elicited in an animal host by immunization with ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin receptor, or β -adrenoceptor-derived immunogenic components or may be formed by *in vitro* immunization of immune cells. The immunogenic components used to elicit the antibodies may be isolated from human
5 cells or produced in recombinant systems. The antibodies may also be produced in recombinant systems programmed with appropriate antibody-encoding DNA.

Alternatively, the antibodies may be constructed by biochemical reconstitution of purified heavy and light chains. The antibodies include hybrid antibodies (*i.e.*, containing two sets of heavy chain/light chain combinations, each of which recognizes a different antigen),
10 chimeric antibodies (*i.e.*, in which either the heavy chains, light chains, or both, are fusion proteins), and univalent antibodies (*i.e.*, comprised of a heavy chain/light chain complex bound to the constant region of a second heavy chain). Also included are Fab fragments, including Fab' and F(ab)₂ fragments of antibodies. Methods for the production of all of the above types of antibodies and derivatives are well-known in the art and are discussed in
15 more detail below. For example, techniques for producing and processing polyclonal antisera are disclosed in Mayer and Walker, *Immunochemical Methods in Cell and Molecular Biology*, 1987 (Academic Press, London). The general methodology for making monoclonal antibodies by hybridomas is well known. Immortal antibody-producing cell lines can be created by cell fusion, and also by other techniques such as
20 direct transformation of B lymphocytes with oncogenic DNA, or transfection with Epstein-Barr virus. See, *e.g.*, Schreier *et al.*, *Hybridoma Techniques*, 1980; U.S. Patent Nos. 4,341,761; 4,399,121; 4,427,783; 4,444,887; 4,466,917; 4,472,500; 4,491,632; and 4,493,890. Panels of monoclonal antibodies produced against ACE, AGT, or AT1-derived epitopes can be screened for various properties; *i.e.* for isotype, epitope affinity, *etc.*

25 The antibodies of this invention can be purified by standard methods, including but not limited to preparative disc-gel electrophoresis, isoelectric focusing, HPLC, reversed-phase HPLC, gel filtration, ion exchange and partition chromatography, and countercurrent distribution. Purification methods for antibodies are disclosed, *e.g.*, in *The Art of Antibody Purification*, 1989 (Amicon Division, W.R. Grace & Co.) General
30 protein purification methods are described in *Protein Purification: Principles and Practice*, R.K. Scopes, Ed., 1987 (Springer-Verlag, New York, NY).

Methods for determining the immunogenic capability of the disclosed sequences and the characteristics of the resulting sequence-specific antibodies and immune

cells are well-known in the art. For example, antibodies elicited in response to a peptide comprising a particular polymorphic sequence can be tested for their ability to specifically recognize that polymorphic sequence, *i.e.*, to bind differentially to a peptide or polypeptide comprising the polymorphic sequence and thus distinguish it from a similar peptide or
5 polypeptide containing a different sequence at the same position.

Diagnostic Methods and Kits

The present invention provides kits for the determination of the sequence at a polymorphic position or positions within the *ACE*, *AT1*, *AGT*, *renin*, *aldosterone*
10 *synthase*, *type-2 angiotensin II receptor*, *endothelin receptor*, or β -*adrenoceptor* gene in an individual, in combination with determination of the sequence at polymorphism positions of other genes. The kits comprise a means for determining the sequence at the polymorphic positions, and may optionally include data for analysis of polymorphic patterns. The means for sequence determination may comprise suitable nucleic acid-based
15 and immunological reagents (see below). Preferably, the kits also comprise suitable buffers, control reagents where appropriate, and directions for determining the sequence at a polymorphic position. The kits may also comprise data for correlation of particular polymorphic patterns with desirable treatment regimens or other indicators.

Nucleic Acid-Based Diagnostic Methods and Kits

The invention provides nucleic acid-based methods for detecting polymorphic patterns in a biological sample. The sequence at particular polymorphic positions in the genes is determined using any suitable means known in the art, including without limitation hybridization with polymorphism-specific probes and direct
25 sequencing.

The present invention also provides kits suitable for nucleic acid-based diagnostic applications. In one embodiment, diagnostic kits include the following components:

- (i) *Probe DNA*: The probe DNA may be pre-labelled; alternatively,
30 the probe DNA may be unlabelled and the ingredients for labelling may be included in the kit in separate containers; and

(ii) *Hybridization reagents*: The kit may also contain other suitably packaged reagents and materials needed for the particular hybridization protocol, including solid-phase matrices, if applicable, and standards.

In another embodiment, diagnostic kits include:

5 (i) *Sequence determination primers*: Sequencing primers may be pre-labelled or may contain an affinity purification or attachment moiety; and

(ii) *Sequence determination reagents*: The kit may also contain other suitably packaged reagents and materials needed for the particular sequencing protocol. In one preferred embodiment, the kit comprises a panel of sequencing primers,
10 whose sequences correspond to sequences adjacent to the polymorphic positions.

Antibody-Based Diagnostic Methods and Kits

The invention also provides antibody-based methods for detecting polymorphic patterns in a biological sample. The methods comprise the steps of: (i)
15 contacting a sample with one or more antibody preparations, wherein each of the antibody preparations is specific for a particular polymorphic form of the gene under conditions in which a stable antigen-antibody complex can form between the antibody and antigenic components in the sample; and (ii) detecting any antigen-antibody complex formed in step (i) using any suitable means known in the art, wherein the detection of a complex indicates
20 the presence of the particular polymorphic form in the sample.

Typically, immunoassays use either a labelled antibody or a labelled antigenic component (*e.g.*, that competes with the antigen in the sample for binding to the antibody). Suitable labels include without limitation enzyme-based, fluorescent, chemiluminescent, radioactive, or dye molecules. Assays that amplify the signals from the
25 probe are also known, such as, for example, those that utilize biotin and avidin, and enzyme-labelled immunoassays, such as ELISA assays.

The present invention also provides kits suitable for antibody-based diagnostic applications. Diagnostic kits typically include one or more of the following components:

30 (i) *Polymorphism-specific antibodies*: The antibodies may be pre-labelled; alternatively, the antibody may be unlabelled and the ingredients for labelling may be included in the kit in separate containers, or a secondary, labelled antibody is provided; and

(ii) *Reaction components*: The kit may also contain other suitably packaged reagents and materials needed for the particular immunoassay protocol, including solid-phase matrices, if applicable, and standards.

The kits referred to above may include instructions for conducting the test.

- 5 Furthermore, in preferred embodiments, the diagnostic kits are adaptable to high-throughput and/or automated operation.

Drug Targets and Screening Methods

According to the present invention, nucleotide sequences derived from the
10 gene encoding a polymorphic form of ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin receptor, or β -adrenoceptor, and peptide sequences derived from that polymorphic for, are useful targets to identify cardiovascular drugs, *i.e.*, compounds that are effective in treating one or more clinical symptoms of cardiovascular disease. Drug targets include without limitation (i) isolated nucleic acids derived from the
15 gene encoding ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin receptor, or β -adrenoceptor and (ii) isolated peptides and polypeptides derived from ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin receptor, or β -adrenoceptor polypeptides, each of which comprises one or more polymorphic positions.

20

In vitro screening methods

In one series of embodiments, an isolated nucleic acid comprising one or more polymorphic positions is tested *in vitro* for its ability to bind test compounds in a sequence-specific manner. The methods comprise:

- 25 (i) providing a first nucleic acid containing a particular sequence at a polymorphic position and a second nucleic acid whose sequence is identical to that of the first nucleic acid except for a different sequence at the same polymorphic position;
- (ii) contacting the nucleic acids with a multiplicity of test compounds under conditions appropriate for binding; and
- 30 (iii) identifying those compounds that bind selectively to either the first or second nucleic acid sequence.

Selective binding as used herein refers to any measurable difference in any parameter of binding, such as, *e.g.*, binding affinity, binding capacity, *etc.*

In another series of embodiments, an isolated peptide or polypeptide comprising one or more polymorphic positions is tested *in vitro* for its ability to bind test compounds in a sequence-specific manner. The screening methods involve:

- (i) providing a first peptide or polypeptide containing a particular sequence at a polymorphic position and a second peptide or polypeptide whose sequence is identical to the first peptide or polypeptide except for a different sequence at the same polymorphic position;
- (ii) contacting the polypeptides with a multiplicity of test compounds under conditions appropriate for binding; and
- (iii) identifying those compounds that bind selectively to one of the nucleic acid sequences.

In preferred embodiments, high-throughput screening protocols are used to survey a large number of test compounds for their ability to bind the genes or peptides disclosed above in a sequence-specific manner.

- Test compounds are screened from large libraries of synthetic or natural compounds. Numerous means are currently used for random and directed synthesis of saccharide, peptide, and nucleic acid based compounds. Synthetic compound libraries are commercially available from Maybridge Chemical Co. (Trevillet, Cornwall, UK), Comgenex (Princeton, NJ), Brandon Associates (Merrimack, NH), and Microsource (New Milford, CT). A rare chemical library is available from Aldrich (Milwaukee, WI). Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available from *e.g.* Pan Laboratories (Bothell, WA) or MycoSearch (NC), or are readily producible. Additionally, natural and synthetically produced libraries and compounds are readily modified through conventional chemical, physical, and biochemical means.

In vivo screening methods

- Intact cells or whole animals expressing polymorphic variants of a gene encoding ACE, AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin receptor, or β -adrenoceptor can be used in screening methods to identify candidate cardiovascular drugs.

In one series of embodiments, a permanent cell line is established from an individual exhibiting a particular polymorphic pattern. Alternatively, cells (including without limitation mammalian, insect, yeast, or bacterial cells) are programmed to express a gene comprising one or more polymorphic sequences by introduction of appropriate

5 DNA. Identification of candidate compounds can be achieved using any suitable assay, including without limitation (i) assays that measure selective binding of test compounds to particular polymorphic variants of the gene; (ii) assays that measure the ability of a test compound to modify (*i.e.*, inhibit or enhance) a measurable activity or function of the gene; and (iii) assays that measure the ability of a compound to modify (*i.e.*, inhibit or

10 enhance) the transcriptional activity of sequences derived from the promoter (*i.e.*, regulatory) regions the gene.

In another series of embodiments, transgenic animals are created in which (i) a human *ACE*, *AT1*, *AGT*, *renin*, *aldosterone synthase*, *type-2 angiotensin II receptor*, *endothelin receptor*, or *β -adrenoceptor* gene having different sequences at particular

15 polymorphic positions are stably inserted into the genome of the transgenic animal; and/or (ii) the endogenous genes are inactivated and replaced with human genes having different sequences at particular polymorphic positions. See, *e.g.*, Coffman, Semin. Nephrol., 1997, 17:404; Esther *et al.*, Lab. Invest., 1996, 74:953; Murakami *et al.*, Blood Press. Suppl., 1996, 2:36. Such animals can be treated with candidate compounds and monitored for one

20 or more clinical markers of cardiovascular status.

Furthermore, populations that are not amenable to an established treatment for a cardiovascular disease or disorder can be selected for testing of alternative treatments. Moreover, treatments that are not as effective in the general population, but that are highly effective in the selected population, may be identified that otherwise would

25 be overlooked. This is an especially powerful advantage of the present invention, since it eliminates some of the randomness associated with clinical trials.

The following are intended as non-limiting examples of the invention.

30 **Example 1: Methods for Identification of Polymorphic Positions in Human Genes Encoding ACE, AGT, and AT1**

The following studies were performed to identify polymorphic residues within the genes encoding human ACE, AGT, and AT1.

DNA samples were obtained from 277 individuals. The individuals were Caucasian males born in Uppsala, Sweden between 1920 and 1924. Individuals were selected for the test population based on their medical history, *i.e.*, they were either (i) healthy, with no signs of cardiovascular disease (100); or (ii) had suffered one of acute
5 myocardial infarction (68), silent myocardial infarction (34), stroke (18), stroke and acute myocardial infarction (19), or high blood pressure at age 50 (39). DNA samples were obtained from each individual.

DNA sequence analysis was carried out by: (i) amplifying short fragments of each of the ACE, AGT, and AT1 genes using polymerase chain reaction (PCR) and (ii)
10 sequencing the amplified fragments. The sequences obtained from each individual were then compared with known ACE, AGT, and AT1 genomic sequences (see Table 1).

(i) Amplification: PCR reactions employed the primers shown in Table 2 below.

Table 2 - PCR Primers

Name	Sequence	SEQ ID NO	Modification *)	Nucleotides	Numbering according to **)
ACE/79RB	5'-TGC GTGCTTCAG AAGTCC-3'	1	B	158-175	i+20: 1-175
ACE/82RB	5'-CCAGG GAGGTGAAGAAATC-3'	2	B	35-53	e20, J04144
ACE/84FT	5'-AGCCAGGCAGTAATGACCT-3'	3	T	1-19	i-19: 1-218
ACE/94FB	5'-GCCCACTGTTCCCTTATG-3'	4	B	1-18	i-21: 1-76
ACE/95RB	5'-TGCCCTGACTGACAGAGC-3'	5	B	105-122	i+23: 1-122
ACE/96RT	5'-GCCCTGGTGTGCCTGT-3'	6	T	1-16	i-22: 1-65
ACE/107F	5'-TGCCTGGATATGTGTTGC-3'	7	-	1-18	i-15: 1-225
ACE/107FB	5'-TGCCTGGATATGTGTTGC-3'	8	B	1-18	i-15: 1-225
ACE/108RB	5'-GCCCTCGCCTCTCACT-3'	9	B	23-38	i+16: 1-38
ACE/111RT	5'-TCCCCTCTCCCTGTACCT-3'	10	T	17-34	i+15: 1-34
ACE/114RB	5'-GTGCTGGGGTAGGGTAGA-3'	11	B	101-118	i+7: 1-118
ACE/118FT	5'-TCCCCTGACCTGGCT-3'	12	T	221-236	i-7: 1-253
ACE/119FB	5'-GGGGCACCGTGATGTT-3'	13	B	1-16	i-4: 1-120
ACE/119FT	5'-GGGGCACCGTGATGTT-3'	14	T	1-16	i-4: 1-120
ACE/120RB	5'-GCCAGAGCCTTTGGTTT-3'	15	B	230-246	i+5: 1-246
ACE/122FB	5'-TGGAAGAGCCGACTTACAG-3'	16	B	1-19	i-5: 1-78
ACE/123RB	5'-TCCCAGAGGCAAGAGG-3'	17	B	225-241	i+4: 1-241
ACE/130F	5'-GTTTCTACTGCGGCTTCAT-3'	18	-	1-19	i-8: 1-131
ACE/130FB	5'-GTTTCTACTGCGGCTTCAT-3'	19	B	1-19	i-8: 1-131

5

10

15

20

Name	Sequence	SEQ ID NO	Modification *)	Nucleotides	Numbering according to **)
ACE/134RB	5'-TCCTGGAAGAGGGAGTTTC-3'	20	B	148-166	i+9: 1-166
ACE/145F	5'-GCAGGATGAGAGCAACAAC-3'	21	-	1-18	i-7: 1-253
ACE/146F	5'-CTGGAGACCACTCCCATCCTTTCT-3'	22	-	1-24	i-17: 1-454
ACE/147R	5'-GATGTGGCCATCACATTCTGTCAGAT-3'	23	-	1-25	e17, J04144
ACE/170RT	5'-CTTCCGTGGGACTCATGT-3'	24	T	23-40	i+5: 1-246
ACE/171RT	5'-TGACCCGTGAGGCTCTA-3'	25	T	136-152	i+8: 1-152
ACE/173F	5'-GCCCAATAGGAGGAAGCA-3'	26	MT	1-10, 1-9	i-2: 1-10, e2
ACE/174R	5'-CCCCCCTCTCCAAAGAA-3'	27	-	166-184	i-2: 1-184
ACE/175FB	5'-GCC-3'		MT, B	1-3	i-2: 1-10
ACE/176RT	5'-TCCCTGATGGGCTGCTCTC-3'	28	T	65-83	i-2: 1-184
ACE/177FT	5'-CAAGGCCCTCAACCACTC-3'	29	T	1-19	i-24: 1-50
ACE/178RB	5'-TTCCCAACAAAAGCTCCAGTG-3'	30	B	71-90	i+24: 1-108
ACE/179R	5'-GGCTCAAAATGGCAAGTGT-3'	31	-	89-108	i+24: 1-108
ACE/180FT	5'-GGGCCATGTCTTCTGACTC-3'	32	T	1-20	i-25: 1-45
ACE/181RB	5'-CAGCCTGGAGGGTTAAGA-3'	33	B	33-51	i+25: 1-51
ACE/182R	5'-CCCTTCTGAGCGAGCTGAGT-3'	34	-	1-6, 1-14	i-26: 1-6, e26, J04144
ACE/183F	5'-GGCCATGTTGAGCTACTTCAA-3'	35	-	83-103	e25, J04144
ACE/184FB	5'-CCTCCAGCCTTGGGTCTTAA-3'	36	B	19-38	i+25: 1-38
ACE/185RT	5'-TTCCCATCCCAGTCTCTGGT-3'	37	T	269-288	e26, J04144
ACE/188RT	5'-GGCAGCCTGTTGATGAGT-3'	38	T	116-134	e17, J04144

5

10

15

20

Name	Sequence	SEQ ID NO	Modification *)	Nucleotides	Numbering according to **)
ACE/192FB	5'-ATTCCAGCTCTGAAATTCTCTGA-3'	39	B	1-23	i-17: 1-85
ACP/3FT	5'-GAGCCCCCTCCAGCACCTC-3'	40	T	499-5017	X94359
ACP/4RB	5'-ACCCGAGCCTGCCACCC-3'	41	B	5302-5318	X94359
ACP/5FT	5'-GGTCGGGCTGGGAAGATC-3'	42	T	5232-5249	X94359
ACP/6RB	5'-TCGGCTCTGCCCCCTTCTC-3'	43	B	5576-5593	X94359 + additional downstream sequence
ACP/7FT	5'-GCCCTTTCTCCAGCTTCCCTCT-3'	44	T	5361-5381	X94359
ACP/8RB	5'-CGGCGGCAGCAGCAACA-3'	45	B	5666-5682	X94359 + additional downstream sequence
ACP/11FB	5'-GAGCCCCCTCCAGCACCTC-3'	46	B	499-5017	X94359
ACP/12RT	5'-ACCCGAGCCTGCCACCC-3'	47	T	5302-5318	X94359
ACP/13FB	5'-GGTCGGGCTGGGAAGATC-3'	48	B	5232-5249	X94359
ACP/14RT	5'-TCGGCTCTGCCCCCTTCTC-3'	49	T	5576-5593	X94359 + additional downstream sequence
ACP/15FB	5'-GCCCTTTCTCCAGCTTCCCTCT-3'	50	B	5361-5381	X94359

Name	Sequence	SEQ ID NO	Modification *)	Nucleotides	Numbering according to **)
ACP/16RT	5'-CGGCGGCAGCAGCAACA-3'	51	T	5666-5682	X94359 + additional downstream sequence
ANG/1FT	5'-ATGGCACTTAAAGGTCAGTTAAT-3'	52	T	336-358	M24686
ANG/2RB	5'-TACGGAAGCCCAAGATT-3'	53	B	726-745	M24686
ANG/5FT	5'-CTCCCCAACGGCTGTCCT-3'	54	T	797-814	M24686
ANG/6RB	5'-AGCAGCAACATCCAGTTCTGT-3'	55	B	1119-1139	M24686
ANG/7FT	5'-TCCACGCTCTCTGGACTT-3'	56	T	1099-1117	M24686
ANG/8RB	5'-CTGATCTCAGCTACACATGGATACTA-3'	57	B	1290-1315	M24686
ANG/15FT	5'-CCTGTCTTGGGTGACTCTTC-3'	58	T	7-26	M24687
ANG/17FB	5'-TTCTGGGCTAAATGGTGACA-3'	59	B	285-304	M24686
ANG/18RT	5'-CTTGTCTTCGGTGTCAAGTTT-3'	60	T	675-695	M24686
ANG/19FB	5'-GGGAGCCTTGGACACACAC-3'	61	B	839-856	M24686
ANG/20RT	5'-AGCCTGCATGAACCTGTCAA-3'	62	T	1147-1167	M24686
ANG/21FB	5'-TGGTGGCGTGTTCACA-3'	63	B	1018-1034	M24686
ANG/22RT	5'-GCCAGAGCCAGCAGAGA-3'	64	T	1264-1280	M24686
ANG/29RB	5'-CCACATCCAGGGGAGAC-3'	65	B	335-352	M24687
ANG/30FB	5'-CCTGTCTTGGGTGACTCTTC-3'	66	B	7-26	M24687
ANG/32RT	5'-CCACATCCAGGGGAGAC-3'	67	T	334-352	M24687
ANP/1FT	5'-GTCCCTTCAGTGCCTTAATAC-3'	68	T	314-334	X15232

Name	Sequence	SEQ ID NO	Modification *)	Nucleotides	Numbering according to **)
ANP/2RB	5'-ACAGCCAGATTGAAAGACACA-3'	69	B	593-613	X15232
ANP/3FT	5'-AACCCCTTTTACTGGTCATGTGA-3'	70	T	492-513	X15232
ANP/4RB	5'-CGCTCATGGGATGTGTGAC-3'	71	B	747-765	X15232
ANP/5FT	5'-TGTTTTCCCCAGTGCTATTAGA-3'	72	T	686-708	X15232
ANP/6RB	5'-GCAGGGTCGAGTTACACATTT-3'	73	B	982-1003	X15232
ANP/7FT	5'-CCTCAGGCTGTCACACACCTA-3'	74	T	909-929	X15232
ANP/8RB	5'-CGGCTTACCTTCTGCTGTAGT-3'	75	B	1246-1266	X15232
ANP/9FB	5'-CTCCTTGAACCTGCTTGTTGTT-3'	76	B	273-293	X15232
ANP/10RT	5'-GCATTGAAAGATGTGCTGTTCT-3'	77	T	548-569	X15232
ANP/11FB	5'-TAACGACTACAAAAGCAAGTCTTAC-3'	78	B	446-469	X15232
ANP/12RT	5'-AGAGGGCAGGGAGAGTCT-3'	79	T	805-823	X15232
ANP/13FB	5'-GGCAGCAGGGTCAGAAAGT-3'	80	B	766-783	X15232
ANP/14RT	5'-GCTGGAGAGGAGGGTTACAT-3'	81	T	1127-1146	X15232
ANP/15FB	5'-TGCAAACTTCGGTAAATGTGT-3'	82	B	970-990	X15232
ANP/16RT	5'-CAGAAACAACGGCAGCTTCT-3'	83	T	1224-1242	X15232
AT1/5FT	5'-ACTGGCTGACTTATGCTTTTACT-3'	84	T	547-570	S77410
AT1/6RB	5'-GGGTTGAATTTTGGGACTCATA-3'	85	B	884-905	S77410
AT1/7FT	5'-GCCAGTTTGCCAGCTATAAT-3'	86	T	809-828	S77410
AT1/8RB	5'-TGATGCCTAGTTGAATCAATACA-3'	87	B	1123-1145	S77410
AT1/9FT	5'-GAAGGCTTATGAAATTCAGAAAGA-3'	88	T	1003-1025	S77410
AT1/10RB	5'-AAAGTCGGTTCAGTCCACATAA-3'	89	B	1535-1556	S77410

Name	Sequence	SEQ ID NO	Modification *)	Nucleotides	Numbering according to **)
AT1/16FB	5'-AAACAGCTTGGTGGTAGTC-3'	90	B	469-490	S77410
AT1/17RT	5'-GCAGGTGACTTTGGCTACAA-3'	91	T	762-781	S77410
AT1/18FB	5'-CCTGTACGCTAGTGTGTTTCTACT-3'	92	B	667-690	S77410
AT1/19RT	5'-AGGAAACAGGAAACCCAGTATAT-3'	93	T	932-955	S77410
AT1/22FB	5'-CTGGATTCCCCACCAAATAT-3'	94	B	1090-1109	S77410
AT1/23RT	5'-TGCTCCTTCTTTCACAAAATTAC-3'	95	T	1438-1460	S77410
ATP/1FT	5'-CTTCCGTTATTATGTGTGATATTAGT-3'	96	T	1244-1269	U07144
ATP/2RB	5'-GCATGTACCTAAAAAGTCCTGTC-3'	97	B	1566-1588	U07144
ATP/5FT	5'-ATTGGCATATCCATCACCTTAA-3'	98	T	1628-1649	U07144
ATP/6RB	5'-GATCTCCCAACTCATGCTATGA-3'	99	B	1961-1982	U07144
ATP/7FT	5'-ATTGGATTCAATTGCGCTACAT-3'	100	T	1846-1867	U07144
ATP/8RB	5'-TTTGGTAATACAGTTGTGGATCATA-3'	101	B	2159-2184	U07144
ATP/9FT	5'-TGCAACTTGGGTAGCATGTC-3'	102	T	2077-2096	U07144
ATP/10RB	5'-AGTCGTCCTCCGTGTCAACTATC-3'	103	B	2370-2390	U07144
ATP/11FB	5'-CGTTGTCTTCCGTTATTATGTGT-3'	104	B	1238-1260	U07144
ATP/12RT	5'-TTATTGCAATGTACCTAAAGTGTA-3'	105	T	1455-1479	U07144
ATP/15FB	5'-GCATTTCATATAAGATCAAAATCAGT-3'	106	B	1600-1624	U07144
ATP/16RT	5'-CACCCCTGATAACAAAACCCAGATA-3'	107	T	1929-1951	U07144
ATP/17FB	5'-CTTCTGGCATCAACCTCACT-3'	108	B	1794-1814	U07144
ATP/18RT	5'-ACTTTTAAGGACGAATTAGAGAACT-3'	109	T	2214-2238	U07144
ATP/19FB	5'-GTCCACCCCTTGAAATTCATAAC-3'	110	B	2115-2136	U07144

Name	Sequence	SEQ ID NO	Modification *)	Nucleotides	Numbering according to **)
ATP/20RT	5'-CCCAACCTCCTCCCTCTC-3'	111	T	2396-2413	U07144
ATP/21FT	5'-GCTCGCTCTCCCTCAGAC-3'	112	T	2310-2328	U07144
ATP/22RB	5'-TCCAGCCGCTCCCCATC-3'	113	B	2657-2673	U07144
ATP/23FB	5'-GCTCGCTCTCCCTCAGAC-3'	114	B	2310-2328	U07144
ATP/24RT	5'-TCCAGCCGCTCCCCATC-3'	115	T	2657-2673	U07144
ATR/1F	5'-GCCCTCAGATAATGTAAGTC-3'	116	-	1353-1374	S77410
ATR/2R	5'-AACCGGCACGAAACTTACT-3'	117	-	1834-1854	S77410
ATR/3aF	5'-GCACCTTCACTACCAAAATGAGCA-3'	118	-	1476-1500	S77410
ATR/4cF	5'-GCACCTTCACTACCAAAATGAGCC-3'	119	-	1476-1500	S77410

Where indicated, the primers were modified in one of the following ways:

(i) a biotin molecule was conjugated to the 5' terminus of the indicated sequence (B); (ii) a sequence of nucleotides derived from M13, 5'-CAGGAAACAGCTATGACT-3' (SEQ ID NO:120), was added at the 5' terminus of the indicated sequence (MT); or (iii) the sequence 5'-AGTCACGACGTTGTAAAACGACGGCCAGT-3' (SEQ ID NO:121) was added at the 5' terminus of the indicated sequence (T = Tail). Nucleotides were numbered according to the GenBank sequences listed in Table 1 where indicated. When the sequences involved were not publicly available, the numbering was as in the following examples: The designation "i-4: 1-200" indicates that the primer sequence is located within the sequence extending 200 bp upstream of, and including, the nucleotide immediately upstream of the first coding nucleotide of exon 4. Similarly, the designation "i+4: 1-200" indicates that the primer sequence is located within the sequence extending from the nucleotide that is located immediately downstream of the last coding nucleotide of exon 4 downstream for 200 bp. In each case, the specific location of the primer sequence is indicated in Table 2 in the column marked "Nucleotides".

The reaction components used for PCR are described in Table 3, and the reaction conditions for PCR are described Table 4, below.

Table 3 -PCR Components

Condition	Components	Volume
A	Ultrapure dNTP Set 2.5 mM (dATP:dCTP:dGTP:dTTP = 1:1:1:1), (Pharmacia Biotech)	4 μ l
	10xPCR buffer II, (Perkin Elmer)	5 μ l
	MgCl ₂ solution 2.5 mM,	3 μ l
	AmpliTaq® DNA polymerase (Perkin Elmer) (5U/ml)	0.15 μ l
	Primer 1	1 μ l
	Primer 2	1 μ l
	DNA solution	1 μ l
	R/O-purified water q.s.	Tot. 50 μ l
B	Ultrapure dNTP Set 2.5 mM (dATP:dCTP:dGTP:dTTP = 2:2:1:1:2), (Pharmacia Biotech)	4 μ l
	10xPCR buffer II, (Perkin Elmer)	5 μ l
	MgCl ₂ solution 2.5 mM, (Perkin Elmer)	3 μ l
	AmpliTaq® DNA polymerase (5U/ml)	0.15 μ l
	Primer 1	1 μ l
	Primer 2	1 μ l
	DNA solution	1 μ l
	R/O-purified water q.s.	Tot. 50 μ l
C	Ultrapure dNTP Set 2.5 mM (dATP:dCTP:dGTP:dTTP = 4:4:1:3:4), (Pharmacia Biotech)	4 μ l
	10xPCR buffer II, (Perkin Elmer)	5 μ l

Condition	Components	Volume
	MgCl ₂ solution 2.5 mM, (Perkin Elmer)	3 µl
	AmpliTaq® DNA polymerase (5U/ml)	0.15 µl
	Primer 1	1 µl
	Primer 2	1 µl
	DNA solution	1 µl
	R/O-purified water q.s.	Tot. 50 µl
D	Ultrapure dNTP Set 2.5 mM (dATP:dCTP:dGTP:dTTP = 6:6:1:5:6), (Pharmacia Biotech)	4 µl
	10xPCR buffer II, (Perkin Elmer)	5 µl
	MgCl ₂ solution 2.5 mM, (Perkin Elmer)	3 µl
	AmpliTaq® DNA polymerase (5U/ml)	0.15 µl
	Primer 1	1 µl
	Primer 2	1 µl
	DNA solution	1 µl
	R/O-purified water q.s.	Tot. 50 µl
E	Ultrapure dNTP Set 2.5 mM (dATP:dCTP:dGTP:dTTP = 4:4:1:3:4), (Pharmacia Biotech)	4 µl
	10xPCR buffer II, (Perkin Elmer)	5 µl
	MgCl ₂ solution 2.5 mM, (Perkin Elmer)	2.5 µl
	DMSO	2.5 µl

Condition	Components	Volume
	AmpliTaqGold® DNA polymerase (5U/ml)	0.5 µl
	Primer 1	1 µl
	Primer 2	1 µl
	DNA solution	1 µl
	R/O-purified water q.s.	Tot. 50 µl
F	Ultrapure dNTP Set 2.5 mM (dATP:dCTP:dGTP:dTTP = 1:1:1:1) (Pharmacia Biotech)	4 µl
	10xPCR buffer II, (Perkin Elmer)	5 µl
	MgCl ₂ solution 2.5 mM, (Perkin Elmer)	2 µl
	AmpliTaq® DNA polymerase (5U/ml)	0.5 µl
	Primer 1	1 µl
	Primer 2	1 µl
	DNA solution	1 µl
	R/O-purified water q.s.	Tot. 50 µl
G	Ultrapure dNTP Set 2.5 mM (dATP:dCTP:dGTP:dTTP = 1:1:1:1) (Pharmacia Biotech)	4 µl
	10xPCR buffer II, (Perkin Elmer)	5 µl
	MgCl ₂ solution 2.5 mM, (Perkin Elmer)	2 µl

Condition	Components	Volume
	AmpliTaq® DNA polymerase (5U/ml)	0.5 µl
	Primer 1	1 µl
	Primer 2	1 µl
	DNA solution	1 µl
	R/O-purified water q.s.	Tot. 50 µl
H	Ultrapure dNTP Set 2.5 mM (dATP:dCTP:dGTP:dTTP = 4:4:1:3:4), (Pharmacia Biotech)	4 µl
	10xPCR buffer II, (Perkin Elmer)	5 µl
	MgCl ₂ solution 2.5 mM, (Perkin Elmer)	4 µl
	AmpliTaqGold® DNA polymerase (5U/ml)	0.5 µl
	Primer 1	1 µl
	Primer 2	1 µl
	DNA solution	1 µl
	R/O-purified water q.s.	Tot. 50 µl

Table 4 - PCR Reaction Conditions

PCR-method	Temperature*)	Time *)	Temperature **)	Time	Temperature	Time	No of cycles ***)
25	94	15 s	55	30 s	72	45 s	35
	72	5 min					1
	22	∞					
27	94	15 s	55	30 s	72	45 s	35
	72	5 min					1
	22	∞					
36	94	2 min					1
	94	15 s	58	30 s	72	45 s	35
	72	5 min					1
	22	∞					
38	94	2 min					1
	94	15 s	60	30 s	72	45 s	15
	72	5 min					1
	22	∞					
40	94	2 min					1
	94	15 s	60	30 s	72	45 s	35
	72	5 min					1
	22	∞					

PCR-method	Temperature*)	Time *)	Temperature **)	Time	Temperature	Time	No of cycles ***)
54	96	5 min					1
	96	30 s	61	30 s	72	45 s	15
	72	5 min					1
	22	∞					
56	96	5 min					1
	96	30 s	61	30 s	72	45 s	35
	72	5 min					1
	22	∞					
64	95	2 min					
	95	15 s	59	30 s	72	45 s	40
	72	5 min					1
	22	∞					
70	95	5 min					
	95	15 s	59	30 s	72	45 s	50
	72	5 min					1
	22	∞					

All temperatures are given in degrees Celsius.

*) indicates the default initial temperatures (°C) and times of the program.

**) indicates the default temperature (°C) of the program.

***) indicates the default number of cycles of the program, referring to the section of the PCR program where three different temperatures are employed.

Any differences are indicated in "Modifications" in Table 5 below.

The amplified fragments are described in Table 5 below with respect to the primers and PCR reaction conditions used for amplification.

Table 5 - Amplified Fragments

Fragment	Primer 1	Primer 2	PCR method	Modifications of PCR method	PCR reaction conditions
ANP1F	ANP/1FT	ANP/2RB	64		A
ANP2F	ANP/3FT	ANP/4RB	64		B
ANP3F	ANP/5FT	ANP/6RB	64	anneal. temp: 48°C	A
ANP4F	ANP/7FT	ANP/8RB	64	anneal. temp: 59°C	D
ANP5R	ANP/9FB	ANP/10RT	64		A
ANP6R	ANP/11FB	ANP/12RT	64		B
ANP7R	ANP/13FB	ANP/14RT	64		A
ANP8R	ANP/15FB	ANP/16RT	64		C
ANGe2f1F	ANG/1FT	ANG/2RB	64		C
ANGe2f3F	ANG/5FT	ANG/6RB	64		C
ANGe2f4F	ANG/7FT	ANG/8RB	64		A
ANGe2f5R	ANG/17FB	ANG/18RT	64		A
ANGe2f7R	ANG/19FB	ANG/20RT	64		A
ANGe2f8R	ANG/21FB	ANG/22RT	64		A
ANGe3F	ANG/15FT	ANG/29RB	64	anneal. temp: 57°C	F
ANGe3R	ANG/30FB	ANG/32RT	64	anneal. temp: 57°C, 45 cycles	A

Fragment	Primer 1	Primer 2	PCR method	Modifications of PCR method	PCR reaction conditions
ACP2F	ACP/3FT	ACP/4RB	70	anneal. temp: 62°C	E
ACP3F	ACP/5FT	ACP/6RB	70	anneal. temp: 58°C	E
ACP4F	ACP/7FT	ACP/8RB	70		E
ACP6R	ACP/11FB	ACP/12RT	70	anneal. temp: 62°C	E
ACP7R	ACP/13FB	ACP/14RT	70	anneal. temp: 58°C	E
ACP8R	ACP/15FB	ACP/16RT	70		E
ACEe2R PCR1	ACE/173F	ACE/174R	38		A
ACEe2R PCR2	ACE/175FB	ACE/176RT	40		A
ACEe4F PCR1	ACE/119FB	ACE/120RB	27		A
ACEe4F PCR2	ACE/119FT	ACE/123RB	25		A
ACEe5R PCR1	ACE/119FB	ACE/120RB	27		A
ACEe5R PCR2	ACE/122FB	ACE/170RT	25		A
ACEe7F PCR1	ACE/145F	ACE/114RB	27		A
ACEe7F PCR2	ACE/118FT	ACE/114RB	25		A
ACEe8R PCR1	ACE/130F	ACE/134RB	27		A
ACEe8R PCR2	ACE/130FB	ACE/171RT	25		A
ACEe15R PCR1	ACE/107F	ACE/108RB	27		A
ACEe15R PCR2	ACE/107FB	ACE/111RT	25		A
ACEe17R	ACE/192FB	ACE/188RT	40	anneal. temp: 63°C, 40 cycles	A
ACEe19F PCR1	ACE/84FT	ACE/79RB	27		A
ACEe19F PCR2	ACE/84FT	ACE/82RB	25		A
ACEe21R PCR1	ACE/94FB	ACE/95RB	27		A
ACEe21R PCR2	ACE/94FB	ACE/96RT	25		A

Fragment	Primer 1	Primer 2	PCR method	Modifications of PCR method	PCR reaction conditions
ACEe24F PCR1	ACE/177FT	ACE/179R	38		A
ACEe24F PCR2	ACE/177FT	ACE/178RB	40		A
ACEe25F PCR1	ACE/180FT	ACE/182R	38		A
ACEe25F PCR2	ACE/180FT	ACE/181RB	40		A
ACEe26R PCR1	ACE/183F	ACE/185RT	54		A
ACEe26R PCR2	ACE/184FB	ACE/185RT	56		A
ACEDI	ACE/146F	ACE/147R	36		A
ATPflF	ATP/1FT	ATP/2RB	64		A
ATPf3F	ATP/5FT	ATP/6RB	64	anneal. temp: 58°C	A
ATPf4F	ATP/7FT	ATP/8RB	64	anneal. temp: 48°C	A
ATPf5F	ATP/9FT	ATP/10RB	64	anneal. temp: 58°C	A
ATPf6R	ATP/11FB	ATP/12RT	64	anneal. temp: 48°C	A
ATPf8R	ATP/15FB	ATP/16RT	64	anneal. temp: 55°C	G
ATPf9R	ATP/17FB	ATP/18RT	64	anneal. temp: 54°C	A
ATPf10R	ATP/19FB	ATP/20RT	64		A
ATPf11F	ATP/21FT	ATP/22RB	64	initial denaturation: 95°C, 12 min.	H
ATPf12R	ATP/23FB	ATP/24RT	64	initial denaturation: 95°C, 12 min.	H
ATle5f2F	AT1/5FT	AT1/6RB	64		A
ATle5f3F	AT1/7FT	AT1/8RB	64		A
ATle5f4F	AT1/9FT	AT1/10RB	64		C
ATle5f6R	AT1/16FB	AT1/17RT	64		A
ATle5f7R	AT1/18FB	AT1/19RT	64		C
ATle5f9R	AT1/22FB	AT1/23RT	64		A

Fragment	Primer 1	Primer 2	PCR method	Modifications of PCR method	PCR reaction conditions
AT1-spec. 1	ATR/1F ATR/3aF	ATR/2R	40	anneal. temp: 63°C	A
AT1-spec. 2	ATR/1F ATR/2R	ATR/4cF	40	anneal. temp: 63°C	A

All of the PCR products (except fragments ACEDI, AT1-spec. 1 and AT1-spec. 2) were subjected to solid phase sequencing according to the protocol commercially available from Pharmacia Biotech. The sequencing reactions are performed with a sequencing primer having a complementary sequence to the "Tail" sequence previously described in Table 2. The nucleotide sequence of the sequencing primer was 5'-CGACGTTGTAAAACGACGGCCAGT-3' (SEQ ID NO:122), and the primer was fluorescently labeled with a Cy-5-molecule on the 5'-nucleotide. The positions carrying a genetic variation were identified by determination of the nucleotide sequence by the use of the ALFexpress™ system commercially available from Pharmacia Biotech.

The detection of the fragment ACEDI was performed by analyzing the sizes of the amplified fragments by gel electrophoresis, where the presence of a shorter PCR product (192 base pairs) indicated the D-allele and a longer PCR product (479 base pairs) indicated the I-allele. The presence of both bands indicated a heterozygote for the two alleles. The detection of the allele-specific reaction of position AT1-1271 was performed by separately running two parallel PCR reactions on the same sample and comparing the sizes of the amplified fragments. A PCR product of 501 base pairs should always be present as a control in both parallel runs, whereas the presence of a PCR product of 378 base pairs in the reaction designated AT1-spec. 1 indicated the presence of an A in this position. The presence of a PCR product of 378 base pairs in the reaction designated AT1-spec. 2 indicated a C in this position. If the shorter PCR product was present in both reactions, the individual is a heterozygote for A and C.

Results

The analysis described above resulted in the identification of polymorphic positions within the regulatory and coding/intron segments of the human genes encoding ACE, AGT, and AT1. The polymorphic positions, the variant nucleotides found at each of the positions, and the PCR fragment in which the polymorphism was identified are shown in Table 6 below. Also shown are the frequencies of each genotype in a population of 90 individuals, expressed as the percent of the study population having that genotype. Polymorphisms that resulted in alternate amino acids in ACE, AGT, or AT1 are also indicated. As used herein below, the designations "AGR", "ACR", and "ATR" refer to the regulatory regions of the human AGT, ACE, and AT1 genes, respectively; and the designations "AGT", "ACE", and "AT1", refer to the coding regions of the AGT, ACE, and AT1 genes.

Table 6 - Polymorphisms Found in ACE, AGT, and AT1 Genes

Gene	Position	Reported genotype	Genetic variation	Frequency (per cent)	Amino acid change	Fragment	Reference (if any)
AGR	395	T	TT-TA-AA	88-11-1	None	ANP11F ANP15R	-
AGR	412	C	CC-CT	99-1	None	ANP11F ANP15R	-
AGR	432	G	GG-GA	81-19	None	ANP11F ANP15R	-
AGR	449	C	TT-TC	92-8	None	ANP11F ANP15R	-
AGR	692	C	CC-CT	81-19	None	ANP12F ANP16R	-
AGR	839	G	GG-GA	93-7	None	ANP13F ANP17R	-
AGR	1007	G	GG-GA	81-19	None	ANP14F ANP17R	-
AGR	1072	G	GG-GA	89-11	None	ANP14F ANP17R	-
AGR	1204	C	CC-CA-AA	67-33	None	ANP14F ANP18R	-
AGR	1218	A	AA-AG-GG	14-55-31	None	ANP14F ANP18R	Inoue, I et. al. J. C. I., 1997, 99: 1786-1789.
AGT	273	C	CC-CT	99-1	None	ANGe21F ANGe25R	-
AGT	620	C	CC-CT	80-20	Thr - Met	ANGe23F ANGe27R	JeunmaîtreX, et al. Cell, 1992, 71:169-180.
AGT	803	T	TT-TC-CC	35-52-13	Met - Thr	ANGe24F ANGe28R	JeunmaîtreX, et al. Cell, 1992, 71:169-180.

Gene	Position	Reported genotype	Genetic variation	Frequency (per cent)	Amino acid change	Fragment	Reference (if any)
AGT	912	C	CC-CT	99-1	None	ANGe3F ANGe3R	-
AGT	997	G	CC	100	Glu - Gln	ANGe3F ANGe3R	-
AGT	1116	G	GG-GA-AA	87-12-1	None	ANGe3F ANGe3R	-
AGT i3	49 Numbering according to GenBank entry M24688	A	AA-AG	80-20	None	ANGe3F ANGe3R	-
AGT	1174	C	CC-CA	99-1	Leu - Met	ANGe4F ANGe4R	-
ACR	5106	C	CC-CT	98-2	None	ACPf2F ACPf6R	-
ACR	5349	A	AA-AT-TT	35-46-19	None	ACPf3F ACPf7R	Villard, E. <i>et al.</i> Am. J. Hum. Genet., 1996, 58: 1268-1278
ACR	5496	T	TT-TC-CC	35-46-19	None	ACPf4F ACPf8R	Villard, E. <i>et al.</i> Am. J. Hum. Genet., 1996, 58: 1268-1278
ACE	375	A	CC	100	None	ACEe2R	-
ACE	582	C	CC-CT	94-6	None	ACEe4F	-
ACE	731	A	AA-AG	96-4	Tyr - Cys	ACEe5R	-
ACE	1060	G	GG-GA	97-3	Gly - Arg	ACEe7F	-
ACE	1215	C	CC-CT-TT	35-42-23	None	ACEe8R	Villard, E. <i>et al.</i> Am. J. Hum. Genet., 1996, 58: 1268-1278
ACE	2193	G	GG-GA-AA	20-57-23	None	ACEe15R	Villard, E. <i>et al.</i> Am. J. Hum. Genet., 1996, 58: 1268-1278

Gene	Position	Reported genotype	Genetic variation	Frequency (per cent)	Amino acid change	Fragment	Reference (if any)
ACE	1451		DD-DI-II	29-54-20	None	ACED1	Villard, E. <i>et al.</i> Am. J. Hum. Genet., 1996, 58: 1268-1278
ACE	2328	A	AA-AG-GG	20-57-23	None	ACEe17R	Villard, E. <i>et al.</i> Am. J. Hum. Genet., 1996, 58: 1268-1278
ACE	2741	G	TT	100	Gly - Val	ACEe19F	-
ACE	3132	C	CC-CT	99-1	None	ACEe21R	-
ACE	3387	T	TT-TC-CC	20-57-23	None	ACEe24F	-
ACE	3503	C	GG	100	Ala - Gly	ACEe25F	-
ACE	3906	G	GG-GA	91-9	None	ACEe26R	-
ATR	1427	A	AA-AT-TT	77-22-1	None	ANP11F ANP16R	-
ATR	1756	T	TT-TA-AA	75-24-1	None	ANP33F ANP18R	-
ATR	1853	T	TT-TG-GG	82-11-1	None	ANP33F ANP18R	-
ATR	2046	T	CC-CT-TT	46-46-8	None	ANP14F ANP19R	-
ATR	2354	A	AA-AC-CC	73-26-1	None	ANP55F ANP10R	-
ATR	2355	G	GG-GC-CC	73-26-1	None	ANP55F ANP10R	-
ATR	2415	A	AA-AG	75-24-1	None	ANP11F ANP12R	-
ATI	449	G	GG-GC	99-1	Ser - Thr	AT1e572F AT1e586R	-
ATI	678	T	CC-CT-TT	31-48-21	None	AT1e533F AT1e577R	Rolfs A, et. al. Eur. Heart. J., 1994, 15: Suppl. D, 108-112.

Gene	Position	Reported genotype	Genetic variation	Frequency (per cent)	Amino acid change	Fragment	Reference (if any)
AT1	1167	A	AA-AG	92-8	None	AT1e5f4F AT1e5f9R	Rolfs A, et. al Eur. Heart. J., 1994, 15: Suppl. D, 108-112.
AT1	1271	A	AA-AC-CC	50-40-10	None	AT1-spec	Bonnardeaux, A. et al. Hypertension, 1994, 24:63-69.

A subset of these polymorphic positions were further analyzed in an additional 187 individuals. Table 7 shows the polymorphic positions, the sequence at these positions, and the genotype frequencies for each position in a population of 277 as described in Example 1 above.

Table 7 - Polymorphic Positions and Frequencies

Gene	Position	Genetic variation	Frequency (per cent)
AGR	395	TT-TA-AA	87-12-7
AGR	432	GG-GA-AA	78-21-1
AGR	449	TT-TC-CC	94-5-1
AGR	692	CC-CT-TT	78-21-1
AGR	839	GG-GA	96-4
AGR	1007	GG-GA-AA	78-21-1
AGR	1072	GG-GA	76-24
AGR	1204	CC-CA-AA	3-27-70
AGR	1218	AA-AG-GG	16-50-34
AGT	620	CC-CT-TT	75-23-2
AGT	803	TT-TC-CC	34-50-16
AGT	1116	GG-GA-AA	83-15-2
ACR	5349	AA-AT-TT	37-44-19
ACR	5496	TT-TC-CC	38-43-19
ACE	1060	GG-GA	96-4
ACE	1215	CC-CT-TT	34-46-20
ACE	2193	GG-GA-AA	22-53-25
ACE	2328	AA-AG-GG	23-52-25
ACE	3387	TT-TC-CC	24-53-23
ACE	3906	GG-GA-AA	86-13-1
ATR	1427	AA-AT-TT	72-26-2
ATR	1756	TT-TA-AA	72-25-3
ATR	1853	TT-TG-GG	73-25-2
ATR	2046	CC-CT-TT	47-41-12
ATR	2354	AA-AC-CC	72-26-2
ATR	2355	GG-GC-CC	71-27-2

Gene	Position	Genetic variation	Frequency (per cent)
ATR	2415	AA-AG-GG	73-25-2
AT1	678	CC-CT-TT	26-51-23
AT1	1167	AA-AG	88-12
AT1	1271	AA-AC-CC	55-36-9

Example 2: Correlation of Polymorphism Patterns with Cardiovascular Disease

The polymorphic positions identified as in Example 1 were correlated with the following markers of cardiovascular status present in the study population: healthy (100 individuals evaluated); myocardial infarction (MI) (120 individuals); stroke (37 individuals); and high blood pressure (BP) (39 individuals). Polymorphic patterns, *i.e.*, combinations of sequences at particular polymorphic positions, that show a statistically significant correlation with one or more of these markers are shown below.

Table 8 - Polymorphism Patterns

ACR 5349 A/T, AGR 1218 A

	Healthy	MI	Stroke	High BP	Total (n)
# of events	3	7	3	5	17
% within group	3	5.8	8.1	12.8	

ACR 5496 C, AGR 1204 A/C

	Healthy	MI	Stroke	High BP	Total (n)
# of events	2	7	3	2	13
% within group	2	5.8	8.1	5.1	

ACR 5496 C/T, AGR 1218 A, AGT 620 C/T

	Healthy	MI	Stroke	High BP	Total (n)
# of events	4	13	1	3	21
% within group	4	10.8	2.7	7.7	

ACE 2193 A, AGR 1204 C, ACE 2328 G

	Healthy	MI	Stroke	High BP	Total (n)
# of events	0	11	3	3	16
% within group	0	9.2	8.1	7.7	

ACE 2193 A, AGR 1204 A/C

	Healthy	MI	Stroke	High BP	Total (n)
# of events	1	1	0	1	3
% within group	1	0.8	0	2.6	

ACE 3387 T, AGR 1218 A

	Healthy	MI	Stroke	High BP	Total (n)
# of events	2	4	1	3	10
% within group	2	3.3	2.7	7.7	

ACE 3387 T, AGT 620 C/T

	Healthy	MI	Stroke	High BP	Total (n)
# of events	1	10	3	2	15
% within group	1	8.3	8.1	5.1	

AGR 1204 A/C, AT1 678 C/T

	Healthy	MI	Stroke	High	Total (n)
# of events	5	23	5	6	37
% within group	5	19.2	13.5	15.4	

AGR 1204 A/C, AT1 1271 A/C

	Healthy	MI	Stroke	High BP	Total (n)
# of events	3	17	3	4	26
% within group	3	14.2	8.1	10.3	

ACE 1215 C, AGR 1204 A/C

	Healthy	MI	Stroke	High BP	Total (n)
# of events	3	13	5	6	25
% within group	3	10.8	13.5	15.4	

AGR 1204 A/C, AT1 1167 A, ACE 3906 A/G

	Healthy	MI	Stroke	High BP	Total (n)
# of events	0	5	1	0	6
% within group	0	4.2	2.7	0	

AGR 1204 A, AGT 620 C, AT1 1271 A, AT1 1167 A, AGR 395 A/T

	Healthy	MI	Stroke	High BP	Total (n)
# of events	1	4	5	3	11
% within group	1	3.3	13.5	7.7	

AGR 1204 A/C, AGT 620 C/T, AT1 1271 A/C, AT1 1167 A, AGR 395 T

	Healthy	MI	Stroke	High BP	Total (n)
# of events	3	13	3	2	20
% within group	3	10.8	8.1	5.1	

AGR 1204 A/C, AGT 620 C/T, AT1 1271 A/C, AT1 1167 A/G, AGR 395 T

	Healthy	MI	Stroke	High BP	Total (n)
# of events	0	2	0	1	3
% within group	0	1.7	0	2.6	

Summary of the three previous polymorphic patterns (which involve the same polymorphic positions):

	Healthy	MI	Stroke	High BP	Total (n)
# of events	4	19	8	6	34
% within group	4	15.8	21.6	15.4	

AGR 1204 A, AT1 678 C, AT1 1167 A, AGR 395 A/T

	Healthy October 12, 1999	MI	Stroke	High BP	Total (n)
# of events	1	2	2	1	5
% within group	1	1.7	5.4	2.6	

AGR 1204 A/C, AT1 678 C/T, AT1 1167 A, AGR 395 T

	Healthy	MI	Stroke	High BP	Total (n)
# of events	3	18	5	4	28
% within group	3	15.0	13.5	10.3	

Summary of the two previous polymorphic patterns:

	Healthy	MI	Stroke	High BP	Total (n)
# of events	4	20	7	5	33
% within group	4	16.7	18.9	12.8	

AGT 620 C/T, AT1 1271 A/C, AT1 1167 A, AGR 395 T

	Healthy	MI	Stroke	High BP	Total (n)
# of events	2	8	1	2	13
% within group	2	6.7	2.7	5.1	

AGT 620 C/T, AT1 1271 A/C, AT1 1167 A/G, AGR 395 T

	Healthy	MI	Stroke	High BP	Total (n)
# of events	0	2	0	1	3
% within group	0	1.7	0	2.6	

AGT 620 C, AT1 1271 A, AT1 1167 A, AGR 395 A/T

	Healthy	MI	Stroke	High BP	Total (n)
# of events	1	4	5	3	11
% within group	1	3.3	13.5	7.7	

Summary of the three previous polymorphic patterns:

	Healthy	MI	Stroke	High BP	Total (n)
# of events	3	14	6	6	27
% within group	3	11.7	16.2	15.4	

AGT 620 C, AT1 678 A, AT1 1167 A, AGR 395 A/T

	Healthy	MI	Stroke	High BP	Total (n)
# of events	1	2	2	1	5
% within group	1	1.7	5.4	2.6	

AGT 620 C/T, AT1 678 C/T; AT1 1167 A, AGR 395 T

	Healthy	MI	Stroke	High BP	Total (n)
# of events	3	15	4	4	24
% within group	3	12.5	10.8	10.3	

Summary of the two previous polymorphic patterns:

	Healthy	MI	Stroke	High BP	Total (n)
# of events	4	17	6	5	29
% within group	4	14.2	16.2	12.9	

ACE 2193 A, AGR 1218 A, AGT 803 A

	Healthy	MI	Stroke	High BP	Total (n)
# of events	2	5	1	3	11
% within group	2	4.2	2.7	7.7	

ACE 2193 A, AGT 620 C/T

	Healthy	MI	Stroke	High BP	Total (n)
# of events	1	11	3	2	16
% within group	1	9.2	8.1	5.1	

ACE 2328 G, AGT 620 C/T

	Healthy	MI	Stroke	High BP	Total (n)
# of events	1	11	3	2	16
% within group	1	9.2	8.1	5.1	

ACE 3387 T, AGR 1204 A/C

	Healthy	MI	Stroke	High BP	Total (n)
# of events	0	10	3	3	15
% within group	0	8.3	8.1	7.7	

Example 3: Correlation Between a Specific Polymorphism Pattern and Treatment Response

The following study was undertaken to define polymorphic patterns in the human ACE, AGT, and/or AT1 genes that predict the efficacy of treatments for cardiovascular disease.

Two groups of hypertensive patients were studied, 41 in the first group and 20 in the second group. The groups were analyzed independently and in combination.

The patients in this population were each treated with one of the following five ACE inhibitors: Captopril, Trandolapril, Lisinopril, Fosinopril, or Enalapril. The effect of the drugs on mean arterial blood pressure was quantified. Mean arterial blood pressure was defined as $\frac{2}{3}$ of the diastolic blood pressure + $\frac{1}{3}$ of systolic blood pressure. The individuals were also categorized as "high responders," *i.e.*, those exhibiting a decrease of more than 16 mm Hg during treatment with an ACE inhibitor drug, and "low responders," *i.e.*, those not exhibiting a decrease of more than 16 mm Hg.

One particular polymorphic pattern, ACE 2193 A/G, AGR 1072 G/G, AT1 1167 A/A, which was present in 51% of the first study population, discriminated between high responders and low responders. In the second group of 20 patients, the pattern was

less prevalent (25%), but the correlation with lowered blood pressure remained statistically significant. Individuals having this polymorphic pattern (designated "1" below) experienced a larger decrease in blood pressure than those lacking this polymorphic pattern (designated "0" below).

5

Table 9 - Response to Treatment Correlates with Polymorphic Pattern

Polymorphic Pattern	Observations	Mean (mm Hg) Change in B.P.	S.D.
0	36	-11.4	8.6
1	25	-18.1	9.7

10

Furthermore, the distribution of high responders and low responders (as defined above) was as follows:

Table 10 - Responder Status Correlates with Polymorphic Pattern

Polymorphic Pattern	Low responder %	High responder %
0	80.1	19.4
1	24.0	76.0

15

Taken together, the results from the two groups indicate that the presence of this polymorphic pattern correlates with an incremental decrease of 6.4-7.3 mm Hg relative to individuals not having this polymorphic pattern.

20

The prevalence of this polymorphic pattern was 41% in this hypertensive population. This suggests that testing for this polymorphic pattern in hypertensive patients, followed by prescribing ACE inhibitors only to those patients having this polymorphic pattern, could increase the response rate from 43% (in a hypertensive population in general) to 76% in hypertensive population selected according to the methods of the invention.

25

If even one polymorphism was absent from the pattern, the high resolution of the variable and therefore the predictive value would be lost.

**Example 4: Correlation Between a Specific Polymorphism Pattern
and Treatment Response or Predisposition to a
Cardiovascular Syndrome**

5 In accordance with the methods disclosed in the above examples, additional correlations of polymorphic patterns with either responsiveness to ACE inhibitors, non-responsiveness to ACE inhibitors, predisposition to myocardial infarction, or predisposition to stroke were derived. These correlations are presented in the following table:

Table 11 - Additional Genetic Signatures Covering Components of the RAAS and Prediction

Genetic Signature	Position 1	Genotype1	Position 2	Genotype2	Position 3	Genotype3	Prediction:
GS1	ACE:2193:	A/G	AGR:1072:	G/G	AT1:1167:	A/A	Response to ACE-inhibitor treatment
GS2	ACE:2193:	A/G	AGR:1072:	G/G	ACE:1060:	G/G	Response to ACE-inhibitor treatment
GS3	ATR:2354:	A/A	AT1:678:	C/T	AT1:1167:	A/A	Response to ACE-inhibitor treatment
GS4	ACR:5496:	C/T	AGR:1204:	A/A	AGR:839:	G/G	Response to ACE-inhibitor treatment
GS5	ACR:5496:	C/T	AGR:1204:	A/A	AGT:620:	C/C	Response to ACE-inhibitor treatment
GS6	ACR:5496:	C/T	ACE:1060:	G/G	AGR:449:	C/T	Response to ACE-inhibitor treatment
GS7	ACR:5496:	C/T	ACE:1215:	C/T	AGR:1204:	A/A	Response to ACE-inhibitor treatment
GS9	ACR:5496:	C/T	ACE:3906:	A/G	AGR:1218:	A/G	Response to ACE-inhibitor treatment
GS10	ACR:5496:	C/T	AGR:449:	C/T	AGR:839:	G/G	Response to ACE-inhibitor treatment
GS11	ACR:5496:	C/T	AGR:449:	C/T	AT1:1271:	A/A	Response to ACE-inhibitor treatment
GS12	ACR:5496:	C/T	AGR:449:	C/T	AGR:1072:	G/G	Response to ACE-inhibitor treatment
GS13	ACR:5496:	C/T	AGR:1072:	G/G	AGR:1204:	A/A	Response to ACE-inhibitor treatment
GS14	ACE:1060:	G/G	AGR:1204:	A/C	AGT:620:	C/C	Non-response to ACE-inhibitor treatment
GS15	ACE:1060:	G/G	AGR:449:	C/T	AGT:620:	C/C	Response to ACE-inhibitor treatment
GS16	ACE:1060:	G/G	AGR:1007:	G/G	AT1:678:	C/T	Response to ACE-inhibitor treatment
GS17	ACE:1215:	C/T	ACE:2193:	A/G	AGR:1204:	A/C	Non-response to ACE-inhibitor treatment
GS18	ACE:1215:	C/T	ACE:3906:	A/G	AGR:1218:	A/G	Response to ACE-inhibitor treatment

Genetic Signature	Position 1	Genotype1	Position 2	Genotype2	Position 3	Genotype3	Prediction:
GS19	ACE:2193:	A/G	ATR:2046:	C/T	ATI:678:	C/C	Non-response to ACE-inhibitor treatment
GS20	AGR:449:	C/T	AGR:1204:	A/A	AGT:620:	C/C	Response to ACE-inhibitor treatment
GS21	AGR:449:	C/T	AGR:1204:	A/A	AGT:1116:	G/G	Response to ACE-inhibitor treatment
GS22	AGR:449:	C/T	AGR:839:	G/G	AGT:620:	C/C	Response to ACE-inhibitor treatment
GS23	AGR:449:	C/T	AGT:620:	C/C	AGT:1116:	G/G	Response to ACE-inhibitor treatment
GS24	AGR:449:	C/T	AGT:1116:	G/G	ATI:1271:	A/A	Response to ACE-inhibitor treatment
GS25	AGR:449:	C/T	AGR:1072:	G/G	AGT:620:	C/C	Response to ACE-inhibitor treatment
GS26	AGR:1007:	G/C	AGR:1072:	G/G	ATI:678:	C/T	Response to ACE-inhibitor treatment
GS27	AGR:1072:	G/G	AGR:1204:	A/A	AGT:1116:	G/G	Response to ACE-inhibitor treatment
GS30	ACR:5496:	C/T	AGR:1204:	A/A	ATI:1167:	A/A	Response to ACE-inhibitor treatment
GS33	ACE:2193:	A/G	ATI:678:	C/C	ATI:1271:	A/C	Non-response to ACE-inhibitor treatment
GS36	AGR:1007:	G/G	AGR:1204:	A/A	AGT:1116:	G/G	Response to ACE-inhibitor treatment
GS37	ACE:2193:	A/A	AGR:1204:	A/C	-	-	Predisposition to MI
GS38	AGR:449:	T/T	AGR:1204:	A/C	ATI:1271:	A/C	Predisposition to MI
GS39	ACE:2193:	A/G	AGR:620:	C/T	AGR:1116:	G/G	Predisposition to MI
GS40	ACE:2193:	A/G	ANP:449:	T/T	ATI:1271:	A/C	Predisposition to MI
GS41	ACR:5349:	A/A	AGR:449:	T/T	ATI:1271:	A/C	Predisposition to MI
GS42	ACE:2193:	A/A	AGR:620:	C/T	ATI:1271:	A/A	Predisposition to MI
GS43	ACE:2193:	A/A	AGR:395:	A/T	-	-	Predisposition to stroke

Genetic Signature	Position 1	Genotype1	Position 2	Genotype2	Position 3	Genotype3	Prediction:
GS44	AGR:1007:	A/G	AGR:1072:	G/G	ATI:1167:	A/A	Predisposition to stroke
GS/45	AGR:395:	A/T	AGR:1072:	G/G	AGR:1218:	A/G	Predisposition to stroke

Polymorphic patterns, *i.e.*, combinations of sequences at particular polymorphic positions, that show a statistically significant correlation with myocardial infarction and stroke are shown below:

Table 12 - Correlation with MI

5

MI (n=120)

Healthy (n=100)

10

15

20

Genetic Signature	Prevalence	MI	Healthy	P-value
ACE:2193 AA ANP:1204 AC	5.0%	100%	0%	0.001
AGR:449 TT AGR:1204 AC AT1:1271 AC	8.2%	89%	11%	0.002
ACE:2193 AG AGT:620 CT AGT:1167 GG	4.5%	100%	0%	0.002
ACE:2193 AG AGR:449 TT AT1:1271 AC	17.7%	74%	26%	0.007
ACR:5349 AA AGR:449 TT AT1:1271 AC	14.5%	75%	25%	0.013
ACE:2193 AA AGT:620 CT AT1:1271 AA	3.6%	100%	0%	0.009

Table 13 - Correlation With Stroke

Stroke (n=37)

Healthy (n=100)

	Genetic Signature	Prevalence	Stroke	Healthy	P-value
5	ACE:2193 AA AGR:395 AT	2.9%	100%	0%	0.005
	AGR:1007 AG AGR:1072 GG AT1:1167 AA	9.5%	85%	15%	0.000008
10	AGR:395 AT AGR:1072 GG AGR:1218 AG	4.4%	83%	17%	0.006

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and the accompanying figures. Such modifications are intended to fall within the scope of the appended claims.

It is further to be understood that all base sizes or amino acid sizes, and all molecular weight or molecular mass values, given for nucleic acids or polypeptides are approximate, and are provided for description.

Various patents, patent applications, methodologies, and publications are cited herein, the disclosures of which are incorporated by reference in their entireties.

Example 5: Correlation Between a Specific Polymorphism Pattern and High Blood Pressure Treatment with a β -Blocker

Using the techniques described in Examples 1 and 2, we investigated polymorphisms in the ADBP1 and ADBR2 genes (regulator/promoter regions and coding regions). The source material was the same as described above.

At the outset, we identified a set of key polymorphisms. Some of these are known; others are new. The are summarized in Table 14.

**Table 14 -Polymorphic Positions in the Genes Encoding
the Beta Adrenergic Receptors 1 and 2**

	Gene:Location	Genotypes		
	B1P:2238	A/A	A/G	G/G
5	B1P:2577	C/C	C/T	T/T
	B1P:2757	A/A	A/G	G/G
	B1R:231	A/A	A/G	G/G
	B1R:758	C/C	C/T	T/T
	B1R:1251	C/C	C/G	G/G
10	B1R:1403	A/A	A/G	G/G
	B1R:1528	A/A	A/C	C/C
	B2P:934	A/A	A/G	G/G
	B2P:987	C/C	C/G	G/G
	B2P:1006	A/A	A/G	G/G
15	B2P:1120	C/C	C/G	G/G
	B2P:1221	C/C	C/T	T/T
	B2P:1541	C/C	C/T	T/T
	B2P:1568	C/C	C/T	T/T
	B2R:839	A/A	A/G	G/G
20	B2R:872	C/C	C/G	G/G
	B2R:1045	A/A	A/G	G/G
	B2R:1284	C/C	C/T	T/T
	B2R:1316	A/A	A/C	C/C
	B2R:1846	C/C	C/G	G/G
25	B2R:2032	A/A	A/G	G/G
	B2R:2068	no insertion	insertion G/G	insertion C/G
	B2R:2070	no insertion	insertion C/C	

Abbreviations used for the denotation of the genes:

- 5 B1P (ADBR1) - putative promoter region of the gene encoding the beta
adrenergic receptor 1
- B1R (ADBR1) - the protein coding region of the gene encoding the beta
adrenergic receptor 1
- B2P (ADBR2) - putative promoter region of the gene encoding the beta
adrenergic receptor 2
- B2R (ADBR2) - the protein coding region of the gene encoding the beta
adrenergic receptor 2

10 These positions, one or a combination of several, can be used to predict a
good response to anti-hypertensive treatment with Beta-blockers.

 Indeed, we identified various polymorphisms/polymorphic patterns that can
be used in the prediction of high response to anti-hypertensive treatment with
beta-blockers. These are shown in Table 15.

Table 15 - Polymorphisms Predictive of β -Blocker Responsitivity

15

GS	Positions - genotypes	Prevalence (%)	P-value	Comments
GS1	B2R:1846 - C/C	7.7	0.0049	Patients show a 8.3 mmHg DBP reduction compared to the group lacking this genotype.
GS2	B2R:1045 - A/A	6.6	0.0105	Patients will show a 8.1 mmHg DBP reduction compared to the group lacking this genotype.
GS3	B2R:1316 - A/A	6.6	0.0105	Patients that will show a 8.1 mmHg DBP reduction compared to the group lacking this genotype.

Not only positions in the beta adrenergic receptors, but also positions in the
renin-angiotensin- aldosterone-system can be used to predict response to anti-hypertensive
treatment with beta- blockers (*see* Table 16). The numbering give according to the
numberings in GenBank as defined in Table 1, *supra*.

Table 16 - RAAS Polymorphism Associated with β -Blocker Responsitivity

GS	Positions - Genotypes	Prevalence (%)	Comments
GS105B	ATR:2046 - C/C + AT1:678 - T/T	10	Patients will show more that 10-mmHg DBP (diastolic blood pressure reduction compared to the group lacking this genotype.

* * *

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and the accompanying figures. Such modifications are intended to fall within the scope of the appended claims.

It is further to be understood that all sizes and all molecular weight or molecular mass values are approximate, and are provided for description.

Patents, patent applications, procedures, and publications cited throughout this application are incorporated herein by reference in their entireties.

WHAT IS CLAIMED IS:

- 1 1. A method for assessing cardiovascular status in an individual to be
2 tested comprising comparing
3 (a) a test polymorphic pattern comprising at least one
4 polymorphic position within a gene encoding a polypeptide selected from
5 the group consisting of renin, aldosterone synthase, type-2 angiotensin II
6 receptor, endothelin receptor, and β -adrenoceptor of the individual, with
7- (b) a reference polymorphic pattern derived from a population
8 of individuals exhibiting a predetermined cardiovascular status; and
9 concluding whether the individual possesses the cardiovascular status based on whether
10 the test pattern matches the reference pattern.
- 1 2. The method according to claim 1, wherein the polymorphic pattern
2 of (a) includes a polymorphic position within a second gene encoding a polypeptide
3 selected from the group consisting of ACE, AT1, and AGT.
- 1 3. The method according to claims 1 or 2, wherein the predetermined
2 cardiovascular status is predisposition to a cardiovascular syndrome.
- 1 4. The method according to claim 3, wherein the cardiovascular
2 syndrome is selected from the group consisting of myocardial infarction, unstable angina,
3 hypertension, atherosclerosis, and stroke.
- 1 5. The method according to claim 3, wherein the reference pattern
2 comprises at least two polymorphisms.
- 1 6. The method according to claim 5, wherein the reference pattern
2 comprises at least three polymorphisms.
- 1 7. The method according to claims 1 or 2, wherein the predetermined
2 cardiovascular status is a response to a cardiovascular treatment regimen.

1 8. The method according to claim 7, wherein the treatment regimen
2 comprises administering a cardiovascular drug selected from the group consisting of an
3 ACE inhibitor, an angiotensin II receptor antagonist, a diuretic, an alpha-adrenoreceptor
4 antagonist, a cardiac glycoside, a phosphodiesterase inhibitor, a beta-adrenoreceptor
5 antagonist, a calcium channel blocker, a HMG-CoA reductase inhibitor, an imidizoline
6 receptor blocker, an endothelin receptor blocker, and an organic nitrite.

1 9. The method according to claim 1, wherein the polymorphic position
2 of the gene is selected from the group consisting of a position in the renin coding region
3 comprising a cytosine to thymine transition in exon 10 that creates a premature stop codon
4 at position 387 resulting in a truncated form of renin with 20 amino acids deleted from the
5 carboxyl terminus; a position in a promoter region of aldosterone synthase numbered -344
6 (with the initiation codon starting at 1); positions in the regulatory region of β -adrenergic
7 receptor-1 gene (designated B1P) numbered 2238, 2440, 2493, 2502, 2577, 2585, 2693,
8 2724, and 2757; positions in the β -adrenergic receptor-1 coding region (designated B1R)
9 numbered 231, 758, 1037, 1251, 1403, and 1528; positions in the β -adrenergic receptor-2
10 gene regulatory region (designated B2P) numbered 932, 934, 987, 1006, 1120, 1221,
11 1541, and 1568; positions in the β -adrenergic receptor-2 gene coding region (designated
12 B2R) numbered 839, 872, 1045, 1284, 1316, 1846, 1891, 2032, 2068, and 2070; positions
13 in the endothelin receptor type A coding region (designated ET_A) numbered 969, 1005,
14 1146, and 2485; and combinations of any of the foregoing.

1 10. The method according to claim 2, wherein the polymorphic position
2 of the second gene is selected from the group consisting of positions in the ACE
3 regulatory region numbered 5106, 5349, and 5496; positions in the ACE coding region
4 numbered 375, 582, 731, 1060, 1215, 2193, 2328, 2741, 3132, 3387, 3503, and 3906;
5 position 1451 in the ACE gene as numbered in GenBank entry X62855; positions in the
6 AGT regulatory region numbered 395, 412, 432, 449, 692, 839, 1007, 1072, 1204, and
7 1218; positions in the AGT coding region numbered 273, 620, 803, 912, 997, 1116, and
8 1174; position 49 in the AGT gene as numbered in GenBank entry M24688; positions in
9 the AT1 regulatory region numbered 1427, 1756, 1853, 2046, 2354, 2355, and 2415; and
10 positions in the AT1 coding region numbered 449, 678, 1167, and 1271.

1 11. An isolated nucleic acid encoding a polypeptide selected from the
2 group consisting of renin, aldosterone synthase, type-2 angiotensin II receptor, endothelin
3 receptor, and β -adrenoceptor in an individual, wherein the nucleic acid comprises a
4 polymorphic position and wherein the polymorphic position is a position in a polymorphic
5 pattern derived from a population of individuals exhibiting a predetermined cardiovascular
6 status.

1 12. A probe which hybridizes at high stringency to a polymorphic
2 position as defined in claim 11.

1 13. A library of nucleic acids comprising a nucleic acid which
2 comprises
3 (a) a polymorphic position of a gene encoding a polypeptide
4 selected from the group consisting of renin, aldosterone synthase, type-2
5 angiotensin II receptor, endothelin receptor, and β -adrenoceptor, wherein
6 the polymorphic position is a position in a polymorphic pattern derived
7 from a population of individuals exhibiting a predetermined cardiovascular
8 status, and
9 (b) nucleic acids comprising other polymorphic positions of
10 human genes that are indicative of cardiovascular status of an individual,
11 wherein the human genes are selected from the group consisting of ACE,
12 AT1, AGT, renin, aldosterone synthase, type-2 angiotensin II receptor,
13 endothelin receptor, and β -adrenoceptor.

1 14. The library of claim 13 wherein the nucleic acids are
2 oligonucleotides.

1 15. A kit for assessing cardiovascular status comprising
2 (a) sequence determination primers and
3 (b) sequence determination reagents,
4 wherein the primers are selected from the group consisting of primers that hybridize to or
5 immediately adjacent to

- 1 (i) a polymorphic position in a gene encoding a
2 polypeptide selected from the group consisting of renin, aldosterone
3 synthase, type-2 angiotensin II receptor, endothelin receptor, and β -
4 adrenoceptor, and
5 (ii) a second polymorphic position in a gene encoding a
6 polypeptide selected from the group consisting of ACE, AT1, AGT,
7 renin, aldosterone synthase, type-2 angiotensin II receptor,
8 endothelin receptor, and β -adrenoceptor.

1 16. The kit of claim 15, wherein the second polymorphic position is
2 selected from the group consisting of positions in the ACE regulatory region numbered
3 5106, 5349, and 5496; positions in the ACE coding region numbered 375, 582, 731, 1060,
4 1215, 2193, 2328, 2741, 3132, 3387, 3503, and 3906; position 1451 in the ACE gene as
5 numbered in GenBank entry X62855; positions in the AGT regulatory region numbered
6 395, 412, 432, 449, 692, 839, 1007, 1072, 1204, and 1218; positions in the AGT coding
7 region numbered 273, 620, 803, 912, 997, 1116, and 1174; position 49 in the AGT gene as
8 numbered in GenBank entry M24688; positions in the AT1 regulatory region numbered
9 1427, 1756, 1853, 2046, 2354, 2355, and 2415; positions in the AT1 coding region
10 numbered 449, 678, 1167, and 1271; a position in the renin coding region comprising a
11 cytosine to thymine transition in exon 10 that creates a premature stop codon at position
12 387 resulting in a truncated form of renin with 20 amino acids deleted from the carboxyl
13 terminus; a position in a promoter region of aldosterone synthase numbered -344 (with the
14 initiation codon starting at 1); positions in the regulatory region of β -adrenergic receptor-1
15 gene (designated B1P) numbered 2238, 2440, 2493, 2502, 2577, 2585, 2693, 2724, and
16 2757; positions in the β -adrenergic receptor-1 gene coding region (designated B1R)
17 numbered 231, 758, 1037, 1251, 1403, and 1528; positions in the β -adrenergic receptor-2
18 regulatory region (designated B2P) numbered 932, 934, 987, 1006, 1120, 1221, 1541, and
19 1568; positions in the β -adrenergic receptor-2 gene coding region (designated B2R)
20 numbered 839, 872, 1045, 1284, 1316, 1846, 1891, 2032, 2068, and 2070; positions in the
21 endothelin receptor type A coding region (designated ET_A) numbered 969, 1005, 1146,
22 and 2485; and combinations of any of the foregoing.

SEQUENCE LISTING

<110> Eurona Medical AB
 Norberg, Leif Torbjorn
 Andersson, Maria Kristina
 Lindstrom, Per Harry Rutger

<120> GENES FOR ASSESSING CARDIOVASCULAR
 STATUS AND COMPOSITONS FOR USE THEREOF

<130> 1248/1E973-US1

<140> TBA
 <141>

<160> 140

<170> FastSEQ for Windows Version 3.0

<210> 1
 <211> 18
 <212> DNA
 <213> "Artificial Sequence"

<400> 1
 tgcgtgcttc agaagtcc 18

<210> 2
 <211> 19
 <212> DNA
 <213> "Artificial Sequence"

<400> 2
 ccaggagggt gaagaaatc 19

<210> 3
 <211> 19
 <212> DNA
 <213> "Artificial Sequence"

<400> 3
 agccaggcag taatgacct 19

<210> 4
 <211> 18
 <212> DNA
 <213> "Artificial Sequence"

<400> 4
 gcccactgtt cccttatg 18

<210> 5
 <211> 18

```

        <212> DNA
        <213> "Artificial Sequence"

        <400> 5
tgccctgact gacagagc                                     18

        <210> 6
        <211> 16
        <212> DNA
        <213> "Artificial Sequence"

        <400> 6
gccctgggtgt gcctgt                                     16

        <210> 7
        <211> 18
        <212> DNA
        <213> "Artificial Sequence"

        <400> 7
tgcctggata tgtgttgc                                     18

        <210> 8
        <211> 18
        <212> DNA
        <213> "Artificial Sequence"

        <400> 8
tgcctggata tgtgttgc                                     18

        <210> 9
        <211> 16
        <212> DNA
        <213> "Artificial Sequence"

        <400> 9
gccctcgct ctcact                                       16

        <210> 10
        <211> 18
        <212> DNA
        <213> "Artificial Sequence"

        <400> 10
tcccctctcc ctgtacct                                    18

        <210> 11
        <211> 18
        <212> DNA
        <213> "Artificial Sequence"

        <400> 11
gtgctggggt agggtaga                                    18

        <210> 12

```

```

      <211> 16
      <212> DNA
      <213> "Artificial Sequence"

      <400> 12
tccccctgac ctggct                                     16

      <210> 13
      <211> 16
      <212> DNA
      <213> "Artificial Sequence"

      <400> 13
ggggcaccgt gatggt                                     16

      <210> 14
      <211> 16
      <212> DNA
      <213> "Artificial Sequence"

      <400> 14
ggggcaccgt gatggt                                     16

      <210> 15
      <211> 17
      <212> DNA
      <213> "Artificial Sequence"

      <400> 15
gccagagcct ttggttt                                    17

      <210> 16
      <211> 19
      <212> DNA
      <213> "Artificial Sequence"

      <400> 16
tggaagagcc gacttacag                                  19

      <210> 17
      <211> 17
      <212> DNA
      <213> "Artificial Sequence"

      <400> 17
tcccagaggc aaagagg                                    17

      <210> 18
      <211> 19
      <212> DNA
      <213> "Artificial Sequence"

      <400> 18
gtttctactg cggcttcac                                  19

```



```

        <210> 19
        <211> 19
        <212> DNA
        <213> "Artificial Sequence"

        <400> 19
g t t t c t a c t g   c g g c t t c a t
                                     19

        <210> 20
        <211> 19
        <212> DNA
        <213> "Artificial Sequence"

        <400> 20
t c c t g g a a g a   g g g a g t t t c
                                     19

        <210> 21
        <211> 19
        <212> DNA
        <213> "Artificial Sequence"

        <400> 21
g c a g g a t g a g   a g c a a c a a c
                                     19

        <210> 22
        <211> 24
        <212> DNA
        <213> "Artificial Sequence"

        <400> 22
c t g g a g a c c a   c t c c c a t c c t   t t c t
                                     24

        <210> 23
        <211> 25
        <212> DNA
        <213> "Artificial Sequence"

        <400> 23
g a t g t g g c c a   t c a c a t t c g t   c a g a t
                                     25

        <210> 24
        <211> 18
        <212> DNA
        <213> "Artificial Sequence"

        <400> 24
c t t c c g t g g g   a c t c a t g t
                                     18

        <210> 25
        <211> 17
        <212> DNA
        <213> "Artificial Sequence"

        <400> 25
t g c a c c g t g a   g g c t c t a
                                     17

```

```

      <210> 26
      <211> 18
      <212> DNA
      <213> "Artificial Sequence"

      <400> 26
gcccaatagg aggaagca                                18

      <210> 27
      <211> 19
      <212> DNA
      <213> "Artificial Sequence"

      <400> 27
cccaccccat ctccaagaa                                19

      <210> 28
      <211> 19
      <212> DNA
      <213> "Artificial Sequence"

      <400> 28
tccctgatgg gctgctctc                                19

      <210> 29
      <211> 19
      <212> DNA
      <213> "Artificial Sequence"

      <400> 29
caaggccctc aaccaactc                                19

      <210> 30
      <211> 20
      <212> DNA
      <213> "Artificial Sequence"

      <400> 30
ttcccacaaa agtccagtg                                20

      <210> 31
      <211> 20
      <212> DNA
      <213> "Artificial Sequence"

      <400> 31
ggctcaaaat ggcaagtgtt                                20

      <210> 32
      <211> 20
      <212> DNA
      <213> "Artificial Sequence"

      <400> 32
gggccatgtc cttctgactc                                20

```

<210> 33
<211> 19
<212> DNA
<213> "Artificial Sequence"

<400> 33
cagcctggag gggtaaga 19

<210> 34
<211> 20
<212> DNA
<213> "Artificial Sequence"

<400> 34
cccttctgag cgagctgagt 20

<210> 35
<211> 21
<212> DNA
<213> "Artificial Sequence"

<400> 35
ggccatgtg agctacttca a 21

<210> 36
<211> 20
<212> DNA
<213> "Artificial Sequence"

<400> 36
cctccagcct tgggtcttaa 20

<210> 37
<211> 20
<212> DNA
<213> "Artificial Sequence"

<400> 37
ttcccatccc agtctctggt 20

<210> 38
<211> 19
<212> DNA
<213> "Artificial Sequence"

<400> 38
ggcagcctgg ttgatgagt 19

<210> 39
<211> 23
<212> DNA
<213> "Artificial Sequence"

<400> 39
attccagctc tgaaattctc tga 23

<210> 40
<211> 18
<212> DNA
<213> "Artificial Sequence"

<400> 40
gagcccctcc agcacctc 18

<210> 41
<211> 17
<212> DNA
<213> "Artificial Sequence"

<400> 41
acccgagcct gcccacc 17

<210> 42
<211> 18
<212> DNA
<213> "Artificial Sequence"

<400> 42
ggtcgggctg ggaagatc 18

<210> 43
<211> 18
<212> DNA
<213> "Artificial Sequence"

<400> 43
tcggctctgc cccttctc 18

<210> 44
<211> 21
<212> DNA
<213> "Artificial Sequence"

<400> 44
gccctttctc cagcttcctc t 21

<210> 45
<211> 17
<212> DNA
<213> "Artificial Sequence"

<400> 45
cggcggcagc agcaaca 17

<210> 46
<211> 18
<212> DNA
<213> "Artificial Sequence"

<400> 46
gagcccctcc agcacctc 18

<210> 47
 <211> 17
 <212> DNA
 <213> "Artificial Sequence"

 <400> 47
 acccgagcct gccacc 17

 <210> 48
 <211> 18
 <212> DNA
 <213> "Artificial Sequence"

 <400> 48
 ggtcgggctg ggaagatc 18

 <210> 49
 <211> 18
 <212> DNA
 <213> "Artificial Sequence"

 <400> 49
 tcggctctgc cccttctc 18

 <210> 50
 <211> 21
 <212> DNA
 <213> "Artificial Sequence"

 <400> 50
 gccctttctc cagcttcctc t 21

 <210> 51
 <211> 17
 <212> DNA
 <213> "Artificial Sequence"

 <400> 51
 cggcggcagc agcaaca 17

 <210> 52
 <211> 23
 <212> DNA
 <213> "Artificial Sequence"

 <400> 52
 atggcactta aaggtcagtt aat 23

 <210> 53
 <211> 19
 <212> DNA
 <213> "Artificial Sequence"

 <400> 53
 tacggaagcc caagaagtt 19

```

      <210> 54
      <211> 18
      <212> DNA
      <213> "Artificial Sequence"

      <400> 54
ctccccaacg gctgtctt                                     18

      <210> 55
      <211> 21
      <212> DNA
      <213> "Artificial Sequence"

      <400> 55
agcagcaaca tccagttctg t                                 21

      <210> 56
      <211> 19
      <212> DNA
      <213> "Artificial Sequence"

      <400> 56
tcccacgctc tctggactt                                     19

      <210> 57
      <211> 26
      <212> DNA
      <213> "Artificial Sequence"

      <400> 57
ctgatctcag ctacacatgg atacta                             26

      <210> 58
      <211> 20
      <212> DNA
      <213> "Artificial Sequence"

      <400> 58
cctgtcttgg gtgactcttc                                     20

      <210> 59
      <211> 20
      <212> DNA
      <213> "Artificial Sequence"

      <400> 59
ttctgggcta aatgggtgaca                                   20

      <210> 60
      <211> 21
      <212> DNA
      <213> "Artificial Sequence"

      <400> 60
cttgtcttcg gtgtcaagtt t                                 21

```

<210> 61
 <211> 18
 <212> DNA
 <213> "Artificial Sequence"

 <400> 61
 gggagccttg gaccacac 18

 <210> 62
 <211> 20
 <212> DNA
 <213> "Artificial Sequence"

 <400> 62
 agcctgcatg aacctgtcaa 20

 <210> 63
 <211> 17
 <212> DNA
 <213> "Artificial Sequence"

 <400> 63
 tggtagggcgt gttcaca 17

 <210> 64
 <211> 17
 <212> DNA
 <213> "Artificial Sequence"

 <400> 64
 gccagagcca gcagaga 17

 <210> 65
 <211> 18
 <212> DNA
 <213> "Artificial Sequence"

 <400> 65
 ccacattcca ggggagac 18

 <210> 66
 <211> 20
 <212> DNA
 <213> "Artificial Sequence"

 <400> 66
 cctgtcttgg gtgactcttc 20

 <210> 67
 <211> 18
 <212> DNA
 <213> "Artificial Sequence"

 <400> 67
 ccacattcca ggggagac 18

<210> 68
 <211> 21
 <212> DNA
 <213> "Artificial Sequence"
 <400> 68
 gtcccttcag tgcctaata c 21
 <210> 69
 <211> 21
 <212> DNA
 <213> "Artificial Sequence"
 <400> 69
 acagccagat tgaaagacac a 21
 <210> 70
 <211> 22
 <212> DNA
 <213> "Artificial Sequence"
 <400> 70
 aaccctttta ctggatcatgt ga 22
 <210> 71
 <211> 19
 <212> DNA
 <213> "Artificial Sequence"
 <400> 71
 cgctcatggg atgtgtgac 19
 <210> 72
 <211> 23
 <212> DNA
 <213> "Artificial Sequence"
 <400> 72
 tgttttcccc agtgtctatt aga 23
 <210> 73
 <211> 21
 <212> DNA
 <213> "Artificial Sequence"
 <400> 73
 gcagggtcga gttacacatt t 21
 <210> 74
 <211> 21
 <212> DNA
 <213> "Artificial Sequence"
 <400> 74
 cctcaggctg tcacacacct a 21


```

        <210> 75
        <211> 21
        <212> DNA
        <213> "Artificial Sequence"

        <400> 75
cggccttacct tctgctgtag t                                21

        <210> 76
        <211> 21
        <212> DNA
        <213> "Artificial Sequence"

        <400> 76
ctccttgaac ctgcttgtgt t                                21

        <210> 77
        <211> 22
        <212> DNA
        <213> "Artificial Sequence"

        <400> 77
gcattgaaag atgtgctgtt ct                                22

        <210> 78
        <211> 25
        <212> DNA
        <213> "Artificial Sequence"

        <400> 78
taacgactac aaaagcaagt cttac                            25

        <210> 79
        <211> 19
        <212> DNA
        <213> "Artificial Sequence"

        <400> 79
agagggcagg ggagagtct                                    19

        <210> 80
        <211> 18
        <212> DNA
        <213> "Artificial Sequence"

        <400> 80
ggcagcaggg tcagaagt                                    18

        <210> 81
        <211> 20
        <212> DNA
        <213> "Artificial Sequence"

        <400> 81
gctggagagg agggttacat                                20

```

<210> 82
 <211> 21
 <212> DNA
 <213> "Artificial Sequence"

 <400> 82
 tgcaaacttc ggtaaattgtg t 21

 <210> 83
 <211> 19
 <212> DNA
 <213> "Artificial Sequence"

 <400> 83
 cagaacaacg gcagcttct 19

 <210> 84
 <211> 24
 <212> DNA
 <213> "Artificial Sequence"

 <400> 84
 actggctgac ttatgctttt tact 24

 <210> 85
 <211> 22
 <212> DNA
 <213> "Artificial Sequence"

 <400> 85
 ggggtgaatt ttgggactca ta 22

 <210> 86
 <211> 20
 <212> DNA
 <213> "Artificial Sequence"

 <400> 86
 gccagtttgc cagctataat 20

 <210> 87
 <211> 23
 <212> DNA
 <213> "Artificial Sequence"

 <400> 87
 tgatgcctag ttgaatcaat aca 23

 <210> 88
 <211> 23
 <212> DNA
 <213> "Artificial Sequence"

 <400> 88
 gaaggcttat gaaattcaga aga 23

<210> 89
<211> 22
<212> DNA
<213> "Artificial Sequence"

<400> 89
aaagtcggtt cagtccacat aa 22

<210> 90
<211> 22
<212> DNA
<213> "Artificial Sequence"

<400> 90
aaacagcttg gtggtgatag tc 22

<210> 91
<211> 20
<212> DNA
<213> "Artificial Sequence"

<400> 91
gcaggtgact ttggctacaa 20

<210> 92
<211> 23
<212> DNA
<213> "Artificial Sequence"

<400> 92
aggaaacagg aaaccagta tat 23

<210> 93
<211> 24
<212> DNA
<213> "Artificial Sequence"

<400> 93
cctgtacgct agtgtgtttc tact 24

<210> 94
<211> 20
<212> DNA
<213> "Artificial Sequence"

<400> 94
ctggattccc caccaaatat 20

<210> 95
<211> 23
<212> DNA
<213> "Artificial Sequence"

<400> 95
tgctccttct ttcacaaaat tac 23

```

      <210> 96
      <211> 26
      <212> DNA
      <213> "Artificial Sequence"

      <400> 96
cttccggttat tatgtgtgat attagt                26

      <210> 97
      <211> 23
      <212> DNA
      <213> "Artificial Sequence"

      <400> 97
gcatgtacct aaaaagtcct gtc                    23

      <210> 98
      <211> 22
      <212> DNA
      <213> "Artificial Sequence"

      <400> 98
attggcatat ccatcacctt aa                      22

      <210> 99
      <211> 22
      <212> DNA
      <213> "Artificial Sequence"

      <400> 99
gatctcccaa ctcatgctat ga                      22

      <210> 100
      <211> 22
      <212> DNA
      <213> "Artificial Sequence"

      <400> 100
attggattca attgcctac at                      22

      <210> 101
      <211> 25
      <212> DNA
      <213> "Artificial Sequence"

      <400> 101
tttggaata cagttgtgga tcata                    25

      <210> 102
      <211> 20
      <212> DNA
      <213> "Artificial Sequence"

      <400> 102
tgcaacttgg gtagcatgtc                        20

```

```

      <210> 103
      <211> 21
      <212> DNA
      <213> "Artificial Sequence"

      <400> 103
agtcgtcccg tgtcaactat c                               21

      <210> 104
      <211> 23
      <212> DNA
      <213> "Artificial Sequence"

      <400> 104
cgttgtcttc cgttattatg tgt                               23

      <210> 105
      <211> 25
      <212> DNA
      <213> "Artificial Sequence"

      <400> 105
ttattgcatg tacctaaaaa gtgta                             25

      <210> 106
      <211> 25
      <212> DNA
      <213> "Artificial Sequence"

      <400> 106
gcattcatat aaagatcaaa tcagt                             25

      <210> 107
      <211> 23
      <212> DNA
      <213> "Artificial Sequence"

      <400> 107
caccttgata acaaaaccag ata                               23

      <210> 108
      <211> 21
      <212> DNA
      <213> "Artificial Sequence"

      <400> 108
ctttctggca tcaacctcac t                               21

      <210> 109
      <211> 25
      <212> DNA
      <213> "Artificial Sequence"

      <400> 109
acttttaagg acgaattaga gaact                             25

```

```

<210> 110
<211> 22
<212> DNA
<213> "Artificial Sequence"

<400> 110
gtccaccctt gaatttcata ac                22

<210> 111
<211> 18
<212> DNA
<213> "Artificial Sequence"

<400> 111
cccaaccttc tccctctc                    18

<210> 112
<211> 19
<212> DNA
<213> "Artificial Sequence"

<400> 112
gctcgctctc cctcacgac                    19

<210> 113
<211> 17
<212> DNA
<213> "Artificial Sequence"

<400> 113
tccagccgct ccccatc                      17

<210> 114
<211> 19
<212> DNA
<213> "Artificial Sequence"

<400> 114
gctcgctctc cctcacgac                    19

<210> 115
<211> 17
<212> DNA
<213> "Artificial Sequence"

<400> 115
tccagccgct ccccatc                      17

<210> 116
<211> 22
<212> DNA
<213> "Artificial Sequence"

<400> 116
gcccctcaga taatgtaagc tc                22

```

```

      <210> 117
      <211> 21
      <212> DNA
      <213> "Artificial Sequence"

      <400> 117
aaccggcacg aaaactttac t                               21

      <210> 118
      <211> 22
      <212> DNA
      <213> "Artificial Sequence"

      <400> 118
gcacttcact accaaatgag ca                               22

      <210> 119
      <211> 22
      <212> DNA
      <213> "Artificial Sequence"

      <400> 119
gcacttcact accaaatgag cc                               22

      <210> 120
      <211> 18
      <212> DNA
      <213> "Artificial Sequence"

      <400> 120
caggaaacag ctatgact                                   18

      <210> 121
      <211> 29
      <212> DNA
      <213> "Artificial Sequence"

      <400> 121
agtcacgacg ttgtaaaacg acggccagt                       29

      <210> 122
      <211> 24
      <212> DNA
      <213> "Artificial Sequence"

      <400> 122
cgacgttgta aaacgacggc cagt                             24

      <210> 123
      <211> 1278
      <212> DNA
      <213> Homo sapien

      <400> 123
ccagacaagt gatttttgag gagtccctat ctataggaac aaagtaatta aaaaaatgta 60

```

tttcagaatt	tacaggccca	tgtgagatat	gattttttta	aatgaagatt	tagagtaatg	120
ggtaaaaaag	aggtatttgt	gtgtttgttg	attgttcagt	cagtgaatgt	acagcttctg	180
cctcatatcc	aggcaccatc	tcttcctgct	ctttgttgtt	aaatgttcca	ttcctgggta	240
atttcattgtc	tgccatcgtg	gatatgccgt	ggctccttga	acctgcttgt	gttgaagcag	300
gatcttcctt	cctgtccctt	cagtgcccta	ataccatgta	tttaaggctg	gacacatcac	360
cactcccaac	ctgcttcacc	cactgctca	cttgtgatca	ctggcttctg	gcgactctca	420
ccaaggtctc	tgtcatgccc	tgttataacg	actacaaaag	caagtcttac	ctataggaaa	480
ataagaatta	taaccctttt	actggtcagt	tgaaacttac	catttgcaat	ttgtacagca	540
taaacacaga	acagcacatc	tttcaatgcc	tgcactctga	aggcattttg	tttgtgtctt	600
tcaatctggc	tgtgctattg	ttgggtgttt	acagtctccc	cagctacact	ggaaacttcc	660
agaaggcact	tttcacttgc	tttgtgtgtt	tccccagtgt	ctattagagg	cctttgcaca	720
gggtaggctc	tttggagcag	ctgaaggcca	cacatcccat	gagcgggcag	cagggtcaga	780
agtggccccc	gtgttgccca	agcaagactc	tccccgtgcc	tctgccctct	gcacctccgg	840
cctgcatgtc	cctgtggcct	cttgggggta	catctcccg	ggctgggtca	gaaggcctgg	900
gtggttggcc	tcaggctgtc	acacacctag	ggagatgtc	ccgtttctgg	gaaccttggc	960
cccgaactcc	gcaaacttgc	gtaaatgtgt	aactcgacc	tgcaccggct	cactctgttc	1020
agcagtga	ctctgcatcg	atcactaaga	cttcctggaa	gaggtcccag	cgtgagtgtc	1080
gcttctggca	tctgtccctc	tggccagcct	gtggctctgg	caagtgatgt	aacctctctc	1140
tccagcctgt	gcacaggcag	cctgggaaca	gtcccatccc	cacctctcag	ctataaatag	1200
ggcctcgtga	cccggccagg	ggaagaagct	gccgttgttc	tgggtactac	agcagaaggt	1260
aagccggggg	ccccctca					1278

<210> 124

<211> 1347

<212> DNA

<213> Homo sapien

<400> 124

gtccatctcc	tcctctcctc	ttctcataag	gacacaggctc	atattagatc	agggctcacc	60
ctcatggcct	cattttaact	taatcatctc	tttaaagatc	ctgtctccaa	ataatgggtca	120
cattctaggt	cctgggggtt	aggacttcaa	cacgggcatt	atggccgttg	gggaggtagg	180
acataattca	gctgatattg	tgcattttgc	acttggatca	tgtagatatt	ttccatggag	240
ctttgaatcc	atttcttctt	ttttttgtag	acatgaatga	tttattctgg	gctaaatggg	300
gacaggaata	ttgagacaat	gaaagatctg	gttagatggc	acttaaagg	cagtttaataa	360
ccacctttca	ccctttgcaa	aatgatattt	caggatgcg	gaagcgagca	ccccagctctg	420
agatggctcc	tgccgggtgtg	agcctgaggg	ccaccatcct	ctgcctcctg	gcctgggctg	480
gcctggctgc	aggtgaccgg	gtgtacatac	accccttcca	cctcgtcatc	cacaatgaga	540
gtacctgtga	gcagctggca	aaggccaatg	ccgggaagcc	caaagacccc	accttcatac	600
ctgctccaat	tcaggccaag	acatccccctg	tggatgaaaa	ggccctacag	gaccagctgg	660
tgctagtcgc	tgcaaaactt	gacaccgaag	acaagttgag	ggccgcaatg	gtcgggatgc	720
tggccaactt	cttgggcttc	cgtatatatg	gcatgcacag	tgagctatgg	ggcgtgggtcc	780
atggggccac	cgtcctctcc	ccaacggctg	tctttggcac	cctggcctct	ctctatctgg	840
gagccttggg	ccacacagct	gacaggctac	aggcaatcct	gggtgttctt	tgggaaggaca	900
agaactgcac	ctcccggctg	gatgcgcaca	aggtcctgtc	tgccctgcag	gctgtacagg	960
gcctgctagt	ggcccagggtc	agggtgata	gccaggccca	gctgctgctg	tccacgggtg	1020
tgggcgtgtt	cacagcccca	ggcctgcacc	tgaagcagcc	gtttgtgcag	ggcctggctc	1080
tctatacccc	tgtggtcctc	ccacgctctc	tggacttcac	agaactggat	gttgctgctg	1140
agaagattga	cagggtcatg	caggctgtga	caggatggaa	gactggctgc	tcctgatgg	1200
gagccagtgt	ggacagcacc	ctggctttca	acacctacgt	ccacttccaa	ggtgaaggcaa	1260
acctctctgc	tggctctggc	cctaggactt	agtatccatg	tgtagctgag	atcagccagt	1320
caggccttgg	agatgggcag	ggggcag				1347

<210> 125

<211> 377

<212> DNA

<213> Homo sapien

<400> 125

cctgcccctg	tcttgggtga	ctcttccttc	cctgtctcct	gtctgatttc	aggaagatg	60
aagggtctt	ccctgctggc	cgagccccag	gagttctggg	tggacaacag	cacctcagt	120
tctgttccca	tgctctctgg	catgggcacc	ttccagcact	ggagtgcac	ccaggacaac	180
ttctcggtga	ctgaagtggc	cttcaactgag	agcgccctgc	tgctgctgat	ccagcctcac	240
tatgcctctg	acctggacaa	ggaggagggt	ctcactttcc	agcaaaactc	cctcaactgg	300
atgaagaaac	tgtctccccg	gtagagccct	cccgggtctc	cctggaatgt	gggagccaca	360
ctctcctgac	ccaggct					377

<210> 126

<211> 273

<212> DNA

<213> Homo sapien

<400> 126

cctctgggag	agccctcact	gtgtggcctg	gagccttcct	aactgtgcat	catctcccca	60
ggaccatcca	cctgaccatg	ccccaactgg	tgctgcaagg	atcttatgac	ctgcaggacc	120
tgctcgccca	ggctgagctg	cccgccattc	tgacacccga	gctgaacctg	caaaaattga	180
gcaatgaccg	catcagggtg	ggggaggtat	ttaccttcct	tgccctacctg	gtccattgca	240
caggtgagca	tgattaagga	aaagagctat	ggg			273

<210> 127

<211> 945

<212> DNA

<213> Homo sapien

<400> 127

agcccaccgc	cggccctcta	gccctcacga	ccctgggtca	cccatgcgcc	ctcagaatga	60
tcttgatcct	gatgtctggt	cctttgcagg	tgctgaacag	catttttttt	gagcttgaag	120
cggatgagag	agagcccaca	gagtctaccc	aacagcttaa	caagcctgag	gtcttggagg	180
tgaccctgaa	ccgcccattc	ctgtttgctg	tgatgatca	aagcgccact	gccctgcact	240
tcctgggccc	cgtggccaac	ccgctgagca	cagcatgagg	ccaggggccc	agaacacagt	300
gcctggcaag	gcctctgccc	ctggccctttg	aggcaaaggc	cagcagcaga	taacaacccc	360
ggacaaatca	gcgatgtgtc	acccccagtc	tcccaccttt	tcttctaattg	agtcgacttt	420
gagctggaaa	gcagccgttt	ctccttggtc	taagtgtgct	gcattggagt	agcagtagaa	480
gcctgcagcg	gcacaaatgc	acctcccagt	ttgctgggtt	tatttttagag	aatgggggtg	540
gggaggcaag	aaccagtgtt	tagcgcggtg	ctactgttcc	aaaaagaatt	ccaaccgacc	600
agcttggttt	tgaacaaaaa	aagtgttccc	ttttcaagtt	gagaacaaaa	attgggtttt	660
aaaattaaag	tatacatttt	tgcatgtgct	tcggtttgta	tttagtgtct	tgaatgtaag	720
aacatgacct	ccgtgtagtg	tctgtaatac	cttagttttt	tccacagatg	cttggtgattt	780
ttgaacaata	cgtgaaagat	gcaagcacct	gaattttctg	ttgaatgcgg	aacaatagct	840
ggttatttct	cccttggtgt	agtaataaac	gtcttgccac	actaagcctc	caaatttact	900
ctttattaga	cgccaacaga	tgtatacatt	cagccagata	gactg		945

<210> 128

<211> 1856

<212> DNA

<213> Homo sapien

<400> 128

gtgagagctc	atgtgcaggc	tgagtgcagc	gcgagggctg	ggactggcat	ggggcccggg	60
gggtgctggg	gagagcacag	agttgggctc	ccctcgctct	tggggtcagc	gtgcccagga	120
aatgcccttt	cttggtttcc	acgagggggg	cttctctgcc	cactgagagc	cggcacctac	180

tccataccat	gccccgatca	gctgccccto	cctcagaacc	gccctctgct	taagggtgtc	240
cactctctcc	tgtcctctct	gcatgcccgc	cctcagagca	gcgggatctc	aaagttatat	300
ttcatgggct	tggactccaa	atggggggaa	ctcggggaca	ctagctcccc	cgggcctcct	360
ttcgtgaccc	tgcccttgac	ttcctcacct	tctctgtctt	tcctgagccc	ctctcccage	420
atgtgactga	taaggaaatt	gagtcacaca	gccctgaaa	gcgccagact	agaacctgag	480
cctctgattc	ctctcacttc	cctccccctac	cctgccactt	cctactggat	agaagtagac	540
agctcttgac	tgtcctcttt	tctccccact	ggctggtoct	tcttagcccc	agccccgtttg	600
aaagagctca	cccccgacac	aaggaccgcg	acacagatac	ctcccagctc	cctctcaacc	660
caccctttcc	agggttggag	aacttgaggc	ataaacattc	ttccatgagg	aatctccacc	720
cagaaatggg	tctttctggc	ccccagccca	gctcccacat	tagaacaatg	acaaatagaa	780
ggggaaatgg	aaaataaaca	ggagaaacgg	ttttcccagg	acaggggtttg	gcctacaagt	840
tgtggatgtg	ggtacccatg	ccaagtgtga	ggggaggctg	gccgggtgtg	gtggctcatg	900
ctctaattccc	agcacttttg	gaggccaagg	tgagtagatc	acttgaggcc	gttaggtttga	960
gaccagcctg	gccaacatgg	tgaaacccca	tctgtactaa	aaatacaaaa	gttagctggg	1020
cgtggtggta	gatgcctgta	gtcccagcta	cttgggaggc	tgaggcatga	gaatcgcttg	1080
agcccagcca	gggcaataca	gcaagacccc	gtctctacaa	ataaaataca	aaaaattagt	1140
tggatgtggt	ggtgcatgcc	tgtagtctta	gctgctaggg	aggctgagat	ggaaggattg	1200
cttgagcctg	ggaggtcaag	gctgcagtga	gccgagatgg	cgccactgca	ctccagcctg	1260
ggcaacagag	tgagaccctg	tctcagaaag	aaaaaagaga	ggagagagac	tggtgccagg	1320
tcaagcacgc	ccctcacagg	actgctgagg	ccctgcagg	gtctgcagca	tggtgccagg	1380
cgggggactc	tgtaagccac	tgctgggagc	cactcccac	ctttctccca	tttctctaga	1440
cctgctgcct	atacagtcac	tttttttttt	tttttgagac	ggagtctcgc	tctgtcgccc	1500
aggctggagt	gcagtggcgg	gatctcggct	cactgcaacg	tccgcctccc	gggttcacgc	1560
cattctcctg	cctcagcctc	ccaagtagct	gggaccacag	cgcccgcac	tacgcccggc	1620
taattttttg	tatttttagt	agagacgggg	tttcaccgtt	ttagccggga	tggtctcgat	1680
ctcctgacct	cgtagtcgcg	ccgcctcggc	ctcccaaagt	gctgggatta	caggcgtagt	1740
acagtcactt	ttatgtggtt	tcgccaattt	tattccagct	ctgaaattct	ctgagctccc	1800
cttacaagca	gaggtgagct	aagggtgga	gctcaagcca	ttcaaccccc	taccag	1856

<210> 129

<211> 4020

<212> DNA

<213> Homo sapien

<400> 129

gccgagcacc	gcgacccgcg	tcatgggggc	cgcctcgggc	cgcggggggc	cggggctgct	60
gctgccgctg	ccgctgctgt	tgctgctgcc	gccgcagccc	gccctggcgt	tggaccccgg	120
gctgcagccc	ggcaactttt	ctgctgacga	ggccggggcg	cagctcttcg	cgagagctta	180
caactccagc	gccgaacagg	tgctgttcca	gagcgtggcc	gccagctggg	cgacagacac	240
caacatcacc	gcggagaatg	caaggcgcca	ggaggaagca	gccctgctca	gccaggagtt	300
tgccggaggc	tggggccaga	aggccaagga	gctgtatgaa	ccgatctggc	agaacttcac	360
ggacccgcag	ctgcccagga	tcatcggagc	tgtgcgaacc	ctgggctctg	ccaacctgcc	420
cctggctaag	cggcagcagt	acaacgcctt	gctaagcaac	atgagcagga	tctactccac	480
cgccaaggtc	tgccctccca	acaagactgc	cactgtctgg	tccctggacc	cagatctcac	540
caacatcctg	gcttcctcgc	gaagctacgc	catgctcctg	tttgcttggg	agggctggca	600
caacgctcgc	ggcatcccgc	tgaaacgcct	gtacgaggat	ttcactgccc	tcagcaatga	660
agcctacaag	caggacggct	tcacagacac	gggggectac	tggcgtctct	ggtacaactc	720
ccccaccttc	gaggacgatc	tggaacacct	ctaccaacag	ctagagcccc	tctacctgaa	780
cctccatgcc	ttcgtccgcc	gcgcactgca	tcgccgatac	ggagacagat	acatcaacct	840
cagggggacc	atccctgctc	atctgctggg	agacatgtgg	gcccagagct	gggaaaacat	900
ctacgacatg	gtggtgcctt	tcccagacaa	gcccacacct	gatgtcacca	gtactatgct	960
gcagcagggc	tggaaacgcca	cgcacatggt	ccgggtggca	gaggagtctt	tcacctccct	1020
ggagctctcc	cccatgcctc	ccgagttctg	ggaagggtcg	atgctggaga	agccggccga	1080
cgggcgggaa	gtggtgtgcc	acgcctcggc	ttgggacttc	tacaacagga	aagacttcag	1140
gatcaagcag	tgcacacggg	tcacgatgga	ccagctctcc	acagtgcacc	atgagatggg	1200

ccatatacag	tactacctgc	agtacaagga	tctgcccgtc	tccctgcgtc	ggggggccaa	1260
ccccggcttc	catgaggcca	ttggggacgt	gctggcgtc	tcggtctcca	ctcctgaaca	1320
tctgcacaaa	atcgccctgc	tggaccgtgt	caccaatgac	acggaaagt	acatcaatta	1380
cttgctaaaa	atggcactgg	aaaaaattgc	cttctctgcc	tttggtact	tgggtggacca	1440
gtggcgctgg	ggggtcttta	gtgggcgtac	cccccttcc	cgctacaact	tcgactgggtg	1500
gtatcttcga	accaagtatc	aggggatctg	tctctctgtt	acccgaaacg	aaacccactt	1560
tgatgctgga	gctaagtttc	atgttccaaa	tgtgacacca	tacatcaggt	actttgtgag	1620
ttttgtcctg	cagttccagt	tccatgaagc	cctgtgcaag	gaggcaggct	atgagggccc	1680
actgcaccag	tgtgacatct	acgggtccac	caaggcaggg	gccaaagctcc	ggaaggtgct	1740
gcaggctggc	tcctccaggc	cctggcagga	ggtgctgaag	gacatggtcg	gcttagatgc	1800
cctggatgcc	cagccgctgc	tcaagtactt	ccagccagtc	acccagtggc	tgcaggagca	1860
gaaccagcag	aacggcgagg	tcctgggctg	gcccagtagc	cagtggcacc	cgccgttgcc	1920
tgacaactac	ccggaggcca	tagacctggt	gactgatgag	gctgaggcca	gcaagtttgt	1980
ggagggaatat	gaccggacat	cccagggtgt	gtggaacgag	tatgccgagg	ccaactggaa	2040
ctacaacacc	aacatcacca	cagagaccag	caagattctg	ctgcagaaga	acatgcaaat	2100
agccaaccac	accctgaagt	acggcaccca	ggccagggaag	tttgatgtga	accagttgca	2160
gaacaccact	atcaagcgga	tcataaagaa	ggttcaggac	ctagaacggg	cagcgctgcc	2220
tgcccaggag	ctggaggagt	acaacaagat	cctgttggtg	atggaaacca	ctacagcgt	2280
ggccactgtg	tgccacccga	atggcagctg	cctgcagctc	gagccagatc	cctagaatgt	2340
gatggccaca	tcgccgaaat	atgaagacct	gttatgggca	tgggagggct	ggcgagataa	2400
ggcggggaga	gccatcctcc	agttttaccc	gaaatacgtg	gaactcatca	accaggctgc	2460
ccggctcaat	ggctatgtag	atgcagggga	ctcgtggagg	tctatgtacg	agacaccatc	2520
cctggagcaa	gacctggagc	ggctcttcca	ggagctgcag	ccactctacc	tcaacctgca	2580
tgccctacgtg	cgccggggcc	tgccacgtca	ctacggggcc	cagcacatca	acctggaggg	2640
gccccattcct	gctcacctgc	tggggaacat	gtgggcgcag	acctgggtcca	acatctatga	2700
cttggtgggtg	cccttccctt	cagccccctc	gatggacacc	acagaggcta	tgctaaagca	2760
gggctggacg	cccaggagga	tgtttaagga	ggctgatgat	ttcttcacct	ccctggggct	2820
gctgcccggtg	cctcctgagt	tctggaacaa	gtcgtatgctg	gagaagccaa	ccgacgggcg	2880
ggaggtgggtc	tgccacgcct	cggcctggga	cttctacaac	ggcaaggact	tccggatcaa	2940
gcagtgcacc	accgtgaact	tggaggacct	ggtggtggcc	caccacgaaa	tgggccacat	3000
ccagtatttc	atgcagtaca	aagacttacc	tgtggccttg	agggagggtg	ccaaccccg	3060
cttccatgag	gccattgggg	acgtgctagc	cctctcagtg	tctacgcccc	agcacctgca	3120
cagtctcaac	ctgctgagca	gtgagggtgg	cagcgacgag	catgacatca	actttctgat	3180
gaagtggcc	cttgacaaga	tcgcctttat	ccccttcagc	tacctcgtcg	atcagtgcg	3240
ctggagggtta	tttgatggaa	gcatacccaa	ggagaactat	aaccaggagt	ggtggagcct	3300
caggctgaag	taccaggggc	tctgcccccc	agtggcccag	actcaagggtg	actttgacct	3360
agggggccaag	ttccacattc	cttctagcgt	gccttacatc	aggtactttg	tcagcttcat	3420
catccagttc	cagttccacg	aggcactgtg	ccaggcagct	ggccacacgg	gccccctgca	3480
caagtgtgac	atctaccagt	ccaaggaggc	cgggcagcgc	ctggcgaccg	ccatgaagct	3540
gggcttcagt	aggccgtggc	cgaagccat	gcagctgatc	acggggccagc	ccaacatgag	3600
cgctcggcc	atggtgagct	acttcaagcc	gctgctggac	tggctccgca	cggagaacga	3660
gctgcatggg	gagaagctgg	gctggccgca	gtacaactgg	acgccgaact	ccgctcgctc	3720
agaagggccc	ctcccagaca	gcggccgcgt	cagcttctctg	ggcctggacc	tggatgcgca	3780
gcaggcccg	gtgggccagt	ggctgctgct	cttctggggc	atcgccctgc	tggtagccac	3840
cctgggcctc	agccagcggc	tcttcagcat	ccgccacgc	agcctccacc	ggcactccca	3900
cgggcccag	ttcggtccg	aggtggagct	gagacactcc	tgaggtgacc	cggtcgggtc	3960
ggccctgccc	aagggcctcc	caccagagac	tgggatggga	acactggtgg	gcagctgagg	4020

<210> 130
 <211> 2720
 <212> DNA
 <213> Homo sapien

<400> 130
 aagcttgcgtg ggggttttgat agagattttg ttttaacctgt agatcatttg aagattaatg 60

ccattgtaac	gatattaaat	ctttcaatcc	aagaacatgg	aatgtcattc	cattttattta	120
ggctacacct	atttcaacaa	ttctttttgt	ttgttttcag	actacaagtt	ttagatcctt	180
ttgttaaatt	tatttcttag	gggttttttt	gttttgtttt	gttttgttgg	ttgggttggt	240
tgttttgaga	tggagtctca	ctctgtcacc	caggctggag	tgtagtggca	caatctcagc	300
tcacagcaac	ctctacctcc	tgggttcaag	cgattattct	gcctcagcct	cctcctcctg	360
agtagctgga	actacaggca	tgccaccacca	cgcctggcct	tttttttttt	tttttctttt	420
gcatttttag	tagagacagg	gtttcacgat	gttggccagc	ctggtctcga	atccctgacc	480
ttgtgattca	cccacctcgg	cctcccaaaag	tgctgagatt	acaggagtga	gccactacac	540
cagggtcattt	cttgatattt	ttactctttt	gatctatagt	aagtaaaatt	gtttttatct	600
ttgaattttt	aaatttttaa	cacagttcaa	atcagtgtgt	ctgatttcat	ctccttctct	660
aacaaaccag	ggtgccagaa	ctgcttcagt	ttctctgcct	tctctttgtc	tatgatgact	720
aatgtatgaa	ggtatctgct	gcatacaact	ttaaaacttca	cattatcctt	atttctcttg	780
accttgacag	atctggcatc	ttttcacctg	gtcgtaaagca	gaaagtcctt	gatctcctta	840
acttttttag	gcattggcagc	atgtgaggca	gggagaggac	acagaccac	acagcaagtg	900
gtgagaagcc	aacagtggaa	ttgttttctt	aattccattt	gttgattggt	tattgctagt	960
gtatagaaat	acaaactgatt	tttgtatatt	gatcttgtat	tctaaaaact	tgctcaactt	1020
gtttcttagt	tctaatagtt	aattaattga	ttccttaggg	ctttttaata	caagatcatg	1080
tcattctaaa	atagaaattg	ttttactttt	tttctaactt	ggatgccatt	tatctttttt	1140
tcttgtccaa	ttgccttcac	tagaaccttt	agtaacaaagt	taaatagaaa	tgggaagact	1200
agacattttg	tcttgttcct	gatcttagac	ataaaaaagt	tgtcttccgt	tattatgtgt	1260
gatattagtt	aagttaagtt	tttcataaat	aaacttcaca	gtttgaggaa	gttcctattc	1320
ctaatttggt	gagtgttagc	atgaaaaagt	gttgaatttt	gtccaagagt	ttttaaaaat	1380
tttttttagac	aatcatgtag	gctttgtcca	ttttttactt	ctttaaaattt	atttttatttg	1440
atacacaata	gatgtacact	tttttaggtac	atgcaataat	ttaatgcccc	tcactataaa	1500
ttcggagctg	cctcctcgcc	gatgattcca	gcgcctgaca	gccaggaccc	caggcagcag	1560
cgagtgcacg	gacttttttag	gtacatgcaa	taattttaatg	cattcatata	aagatcaaat	1620
cagtgcattt	ggcatatcca	tcaccttaaa	tatttgtctt	tttcttcatg	ctagaaacat	1680
tcaagttatt	ttctccttagc	tactctgaaa	tatacaatag	attactgtaa	actacagtca	1740
ccctactcac	ctatctaaca	ttaattgatt	tttggtaaac	taatctaata	ttgctttctg	1800
gcattcaacct	cacttgacca	tgggtgtatag	tccctttcat	atgttatttg	attcaatttg	1860
cctacatttt	gttgagaatt	tttatctata	ctcttaagaa	atattgatct	gtagtctcgt	1920
gatgtcttta	tctggttttg	ttatcaggtt	gataactggc	tcatagcatg	agttgggaga	1980
tcattccttac	tcttctattt	tttggaagag	tttgtgaaga	attgatatta	tttcttcttt	2040
aaatattttat	tgggtttttta	aaatacattt	ttaaaatgca	acttgggtag	catgtccaat	2100
aggaacaaat	gagtgtccac	ccttgaattt	cataaccctc	ggaattaatc	catgtaatct	2160
atgatccaca	actgtattac	caaagtccga	gttactcata	ggaaagagaa	agaagttctc	2220
taattcgctcc	ttaaaagttt	tccaagttca	gaaaaaaaaa	atgttgaaga	acacgaactc	2280
ccgcaggaaa	tgatactcct	gtacccccag	ctcgctctcc	ctcacgaccc	ctcgctaggg	2340
gggggttcggg	accaggtgaa	cgctgatctg	atagttgaca	cgggacgact	gtggcatcat	2400
ccttgctgcc	gtcaatatcc	cgagagggag	gagggttggg	cgggaggggc	tccggggcgg	2460
ggcggaggag	gaggggaatgc	aaaacagagc	ctcgctcccg	gaaccaaga	agcagcaacg	2520
cccctacta	taaattcgga	gctgcctcct	cgccaatgat	tccagcgctc	gacagccagg	2580
acccagggca	gcagcagagt	acaggacgtc	tggaccggcg	cgccgctagc	agctctgctg	2640
ggcgcggcg	gtgatcgatg	gggagcggct	ggagcggacc	cagcgagtga	gggcgcacag	2700
ccgggacgcc	gaggcggcgg					2720

<210> 131
 <211> 480
 <212> DNA
 <213> Homo sapien

gaattctcag	agetggcgaa	acagtcctggt	ccaagcagcc	tctcagcagt	gccttttcagc	60
ctcccctctc	tgagtctttc	caccccttgc	tggtacttta	gtttcttcca	cttcttagac	120
cacgtgtagt	ttcccaattt	ctcttaccca	aatttgctca	cagggaataa	aataaattaa	180

attagccatt	tacaccacag	tgtgaactta	ataacaccaa	caaaagttcc	aaagctctag	240
ggtctcatag	cacctccaga	tccatgatct	cattcggtgt	ttccaacaat	gttttgcacc	300
aaactggaca	catgcttgct	acttcatcat	cctcatcggt	aacattatta	ttattatcat	360
cattttccag	atgaagaaaa	tgaatcaca	gtcaactgac	agtccaaagg	ctccacagct	420
cagaggaggt	aatcatgtg	cttaattcag	aacttttggc	tcccatcact	atgctcttcc	480

<210> 132
 <211> 2268
 <212> DNA
 <213> Homo sapien

<400> 132						
accccaggca	gcagcgagtg	acaggacgtc	tggaccggcg	cgccgctagc	agctctgccg	60
ggccgcggcg	gtgatcgatg	gggagcggct	ggagcggacc	cagcgagtg	ggggcgacag	120
ccgggacgcc	gaggcggcgg	gcgggagacc	cgaccagcg	cagccggccc	tcggcgggac	180
gtgacgcagc	gcccggggcg	cggtttgat	atgtgacaaa	ttgatctaaa	atggctgggt	240
ttttatctga	ataactcact	gatgccatcc	cagaaagtcg	gcaccaggtg	tatttgatat	300
agtgtttgca	acaaattcga	cccaggtgat	caaaatgatt	ctcaactctt	ctactgaaga	360
tggtattaaa	agaatccaag	atgattgtcc	caaagctgga	aggcataatt	acataattgt	420
catgatctct	actttataca	gtatcatctt	tgtggtggga	atatttggaa	acagcttggt	480
ggtgatagtc	atttactttt	atatgaagct	gaagactgtg	gccagtgtgt	ttcttttgaa	540
tttagcactg	gctgacttat	gctttttact	gactttgcca	ctatgggctg	ctacacagc	600
tatggaatac	cgctggccct	ttggcaatta	cctatgtaag	attgcttcag	cagcgtcag	660
tttcaacctg	tacgctagt	tgtttctact	cacgtgtctc	agcattgatc	gatacctggc	720
tattgttccac	ccaatgaagt	cccgccttcg	acgcacaatg	cttgtagcca	aagtcacctg	780
catcatcatt	tggtctgtgg	caggcttggt	cagtttgcca	gctataatcc	atcgaaatgt	840
atttttcatt	gagaacacca	atattacagt	ttgtgctttc	cattatgagt	cccaaaattc	900
aacccttccg	atagggcttg	gcctgaccaa	aaatatactg	ggtttcctgt	ttccttttct	960
gatcattctt	acaagttata	ctcttatttg	gaaggcccta	aagaaggctt	atgaaattca	1020
gaagaacaaa	ccaagaaatg	atgatatttt	taagataatt	atggcaattg	tgcttttctt	1080
tttcttttcc	tgattcccc	accaaataat	cacttttctg	gatgtattga	ttcaactagg	1140
catcatcgt	gactgtagaa	ttgcagatat	tgtggacacg	gccatgccta	tcaccatttg	1200
tatagcttat	tttaacaatt	gcctgaatcc	tcttttttat	ggctttctgg	ggaaaaaatt	1260
taaaagatat	tttctccagc	ttctaaaata	tattccccca	aaagccaaat	cccactcaaa	1320
cctttcaaca	aaaatgagca	cgcttttccta	ccgcccctca	gataatgtaa	gctcatccac	1380
caagaagcct	gcaccatggt	ttgaggttga	gtgacatgtt	cgaaacctgt	ccataaagta	1440
attttttgaa	agaaggagca	agagaacatt	cctctgcagc	acttcaactac	caaatgagca	1500
ttagctactt	ttcagaattg	aaggagaaaa	tgcattatgt	ggactgaacc	gacttttcta	1560
aagctctgaa	caaaagcttt	tcttttcctt	tgcaacaaga	caaagcaaa	ccacattttg	1620
cattagacag	atgacggctg	ctcgaaagac	aatgtcagaa	actcgatgaa	tgtgttgatt	1680
tgagaaattt	tactgacaga	aatgcaatct	ccctagcctg	cttttgcctt	gttatttttt	1740
atttccacat	aaagggtatt	agaatatatt	aaatcggttag	aggagcaaca	ggagatgaga	1800
gttccagatt	gttctgtcca	gtttccaaag	ggcagtaaa	ttttcgtgcc	ggttttcagc	1860
tattagcaac	tgtgctacac	ttgcacctgg	tactgacat	tttgtacaaa	gatatgctaa	1920
gcagtagtcg	tcaagttgca	gatctttttg	tgaaattcaa	cctgtgtctt	atagggtttac	1980
actgccaaaa	caatgcccgt	aagatggctt	atttgtataa	tggtgttact	aaagtccat	2040
ataaaagtta	aactacttgt	aaaggtgctg	cactggctcc	aagtagtagt	gtcctcctag	2100
tatattagtt	tgatttaata	tctgagaagt	gtatatagtt	tggtgtaaaa	agattatata	2160
tcataaagta	tgcttctctg	tttaaaaaaa	gtatatattc	tacacatata	tatatatgta	2220
tatctatctc	tctaaactgc	tgtaatttga	ttaaaatctg	gcaaagtt		2268

<210> 133
 <211> 826
 <212> DNA
 <213> Homo sapien

<400> 133
 gatctaccca ccttggcctc ccaaagtgtc gggacaggtg tgagccacca tgccctggccc 60
 ctctactctt ataattaaac cagctgttgc ttttcctgcc aagaaaccag tcatgaagat 120
 tcacccatgt tctagatggg aaaactgggc tgtagctggg agaggccagt caggggacaaa 180
 gccaaagtta atatatagaa tggagcttcc agggatatagg gggtgggtct gggctaggga 240
 gctggaaacc taggttttac gcttgctcca gttttgatgt tagccctgac agtgctgttt 300
 ctcatcagcc tctgcctgct ccaggggtca cagggccaag ccagatagag ggctgctagc 360
 gtcaatggac acaagattgc tttccacag ctgtccttcc tccagccct ctgctcccca 420
 tccggaaacc tgggtaccct tcacccacct agctctgtcc cgcagtgaga tttattgctg 480
 actgccctgc catctacccc agggtaataa atcagggcag agcagaattg caatcacccc 540
 atgcatggag tgtataaaag gggaaagggt aaggggagcca cagaacctca gtggatctca 600
 gagagagccc cagactgagg gaagcatgga tggatggaga aggatgcctc gctggggact 660
 gctgctgctg ctctggggct cctgtacctt tgggtctccc acagacacca ccacctttaa 720
 acggtaatgt gtaactcagg cagagaaggg gtgggcaggg gtgtagggtc ccacctcccc 780
 aacaccctgg cttttccaca tgcggtgtca ttcagtcctt acgatc 826

<210> 134
 <211> 6910
 <212> DNA
 <213> Homo sapien

<400> 134
 ggatcctgca aggagggata caaattacat acatttgtca aaaccacag catgttgacc 60
 accaggagga gaccccatgt gactccagga ccctgggtga taacaacgta tcgagattcc 120
 tcacatggaa ccagtgcgct cctgtggtgg aggggtgtacc tgtgtcaggg cagggggtac 180
 gtggacattt tctgcagttt ttgatcaatt ttgcaatgaa ctaaatctgt ggtataaaaa 240
 taaagtctat taaaagaatc caaggctccc tctcatctca cgataagata aagtcccat 300
 ccattttact cctctcagcc ctggagaaag gagaggccag gtcccaccac ctccaccag 360
 catggacccc cagtccagac cccacgcctt ttctcagcat cctcagacca gcaggacttg 420
 cagcaatggg gaattaggca cctgacttct ccttcactca cctttggctg ggggcctcca 480
 gccttgacct tcgctctgag agtctcaggg aggtccagag ccagttctcc catgacgtga 540
 tatgtttcca gagcaggttc ctgggtgaga taaaaggatt tgggctgaac aggggtggagg 600
 gagcattgga atggcactca gggcaaaagg agaggtgtgc gtggcagcgc cctggctgtc 660
 cctgc aaagg gcacgggcac tgggcactag agccgctcgg gcccttagga cgtgctgcc 720
 gtttgaagcc atgccccagc atccaggcaa caggtggctg aggtgctgc agatctggag 780
 ggagcagggt tatgagcacc tgcacctgga gatgcaccag accttccagg agctggggcc 840
 cattttcagg taaagccctc cctggccctc gctgggaaca ccagatccc tgcccctgct 900
 gccaggacc ctgccaggca ctcagcactg ccatttccag caggtcccgg cactctgcat 960
 cctttggagg atgggggaagg agtgcagcac atgctggtct gtggtgctgc cagggcaggg 1020
 gatagtgcag agaaaacccc agctcactgc agagagggca ggactcagaa gcactaaagt 1080
 tgaaagggtc cagggagcca gcaggagggc tttagctgtg aagccgctaa tccaggagca 1140
 gggagggttg acaggagaca ctttggattg ggactgcagg gtggggccac gagggacatg 1200
 accccgtcca gcagggcctc ctgcttggcc ccacaggtac aacttgggag gaccacgcat 1260
 ggtgtgtgtg atgctgcccg aggatgtgga gaagctgcaa caggtggaca gcctgcatcc 1320
 ctgcaggatg atcctggagc cctgggtggc ctacagacaa catcgtgggc acaaagtgtg 1380
 cgtgttcttg ttgtaagcgg cgagttggga gctgagagct gggagcaggg tgggcagcct 1440
 ggggtgtagg gggaggcag agaggtagga cccaaaagca catctgccct gggccctgt 1500
 ggtgggcagt gagggtgagc acccgccca gagacggcc atcctgtggg gtccgctctg 1560
 cactgtgggt tggggaagca gggcggtggt ggagaaatgg gcacgggcac ctctgcagag 1620
 aagacgcaga gcaatgagcc cttctgtgta gtgagaacct gctctgcacc aacctcggcg 1680
 gctgctttct ctgcccgtct ggggactgtc cttccatag gtgagaaac tgaggccctg 1740
 agaaggggac ttccactggc ccaggtcaca ggctgagtg tgagcctggt gttcggcggg 1800
 gccgcagcct ccctcagggc gctcagggtc cctgcagtc tggcaaacct tcctgatggg 1860
 gacagtcagg ggcaggaggc aggtggggac gcaggtggct ggtggttccg ttgttctcag 1920
 aagcaaggca caaggtgggg cggttgatgg cactggggag gatgtttcct ggcccgtgga 1980

gaggggtggcg	cctggtcagg	tgggcaggga	gagggctgatg	cttggagtcg	gtcacctgca	2040
gggatgttgt	cattaggacg	ggggaaggac	tggatgagga	tgtcacagtg	gtgacagccc	2100
ccactccatg	gtaggaaggg	aacgctattg	ggaatagtgg	ggtttaggta	aaagggcacc	2160
cgtgggtcgg	ggccttcact	gagggctggcc	tatagatgac	atctgggaga	gagtcaggac	2220
ccaggaaggg	aggtccagga	ggctgggtgc	gcataaatgga	aggaagggga	gcgctcctgt	2280
ctgtgtgtgt	gtcttgcac	tgtgcacatg	ctgtgtgttt	ctctgtacct	gcattgcaca	2340
tgtgtagtgt	gtgcacgtgt	cgtgtgtgaa	tgtatgtgtg	gtgtgtgtgc	acaagtgtct	2400
gtgtgtgtgc	atgtgcagg	gccggcatgg	gtgtagtgtt	tgtgcacaca	tgcacatgcg	2460
tctcttcaca	catgggtgtg	aggtcttgca	tgggcgcacg	tgtgcatgtg	catcttctgc	2520
ctgtcatcac	tgtcaacagc	tcacagcagc	cagctggaca	taaataaagg	agttttgcag	2580
gaatgtggct	gacaggggaa	attcctcccc	accattccct	gggggcatcc	atggagcccc	2640
cacgcactct	ggctgtgggt	aggtatggcat	gaagcacaaa	gcttgggttc	tgtcctgcag	2700
aagatataga	tgtctcacag	agacagcaga	gcagatgccc	cagaggcact	gtgcccagg	2760
cggggaaggg	tggggaaggag	agggcagcca	ggggctctcc	cctcaggaca	ctgtgtgggt	2820
gaggtgggca	aagcttgaca	acaggggtca	cctcctttct	tggagaaaag	ccctaccctg	2880
ttactacagg	gagggcccg	atgggtgagg	tgggtgccga	cttgggtcgc	caggtccccg	2940
gaatgacctc	agttaccctg	tcagcacctg	tgggcagaag	ctaccatctc	atccctgctt	3000
agacctgagt	ggcctttgcc	cagcacctgg	aggccgctct	gagaaaaagg	tgcagctcga	3060
acacaaacag	gcagcttcta	ccagggcccc	cagtcagctc	cctgcaggcc	gattccccct	3120
ggggacaagg	aggatgggat	acgggtcagg	gcctgtgtct	tgtgtggggc	gcctcacaag	3180
ctctgccctg	gcctctgtag	gaatgggcct	gaatggcgct	tcaaccgatt	gcgctggaac	3240
ccagatgtgc	tgtcgcccaa	ggcctgtcag	aggttcctcc	cgtatgggga	tgcagtggcc	3300
agggacttct	cccaggccct	gaagaagaag	gtgctgcaga	acgcccgggg	gagcctgacc	3360
ctggacgtcc	agcccagcat	cttccactac	accatagaag	gtgtggggcca	tgcgggaagg	3420
tccagcccca	gagaccctgg	agtggccagg	gatggggatg	gaggactgaa	gggagtgtgg	3480
ggaggcagcc	aggagccctg	gggctgcctt	gtgctcagca	gtgcatcttc	cccgcagcca	3540
gcaacttagc	tctttttgga	gagcggctgg	gcctgggttg	ccacagcccc	agttctgcca	3600
gcctgaactt	cctccatgcc	ctggaggcca	tgttcaaatc	caccgtccag	ctcatgttca	3660
tgccccaggag	cctgtctcgc	tggatcagcc	ccaaggtgtg	gaaggagcac	tttgaggcct	3720
gggactgcac	cttccagtac	ggtagggcca	gggacccggg	cagtgcctatg	gggaagggac	3780
accatggggg	cccaatttct	ccttctccac	caccagtggt	ggaatggagg	ccacagggag	3840
gggtcgggga	ttcctcacct	tcctgcgggg	gagattgggt	cgaggctggg	gctgggctgg	3900
gctgatccgg	agaatttggg	atgagagcag	ggagatttgg	gtgtcggggc	agtcgggcca	3960
ggaggaggac	actgaaggat	gcttcccagc	accaagatct	gagggctgtc	ccctgctccc	4020
tggacaggtg	acaactgtat	ccagaaaatc	taccaggaac	tggccttcaa	ccgccctcaa	4080
cactacacag	gcacgtgggc	ggagctcctg	ttgaaggcgg	aactgtcact	agaagccatc	4140
aaggccaact	ctatggaaact	cactgcaggg	agcgtggaca	cggtcaggcc	agcaaccagc	4200
cccacccaga	gaggggtgatg	ccaagcctgc	ctcccaggca	ctgcctgcca	atgccacacg	4260
gcacccacgt	tccccatccc	caggctacag	gccccacatt	tctgttgccc	tcagccttcc	4320
ccctcctttg	ttaagggatg	agatttgcag	gggaggggaa	atgtgagctc	cccctcacat	4380
gagactgagt	ttgcagttac	ctgtgtgggg	atccatgctc	caggctggaa	gaaagtgtga	4440
tgaggccctg	gacacacagc	agctctgtcc	ccactggaaa	gctctgggtg	tacaaggaga	4500
aggagggttg	agaggcagct	ggaggactcc	actgggcacc	cttcccagtg	tgcccgggtca	4560
ccttggggcca	gaaatgtaga	tgcacgggag	ggcagggttg	tggggaaagac	agcagcacag	4620
gctccagcca	gtgcagaggg	gcctgtgggt	gcacagtggg	gagaactcaa	tggaaagcaga	4680
gggagctggg	gctccagaac	tccctggatg	atgctgaggt	gtggccccct	gccctaatgg	4740
tggctgtgag	aacccgccct	gaagaggctg	caggggacct	gggccttggg	ggagatgggg	4800
gtcacctttc	cctgaagaag	tcagggaaatc	tggcccaagt	ggcatcaag	gtttcagatc	4860
cgggggtccca	gggctctgtt	tttgctcagg	gcacgtggtg	ctccacccct	cagagggagg	4920
ttgtcctggg	aggggtgtcc	cgggggctga	gtcctcctgt	gcaaggtctg	accctgcaga	4980
catgtgctct	gtagacagcg	tttcccttgc	tgtgacgct	ctttgagctg	gctcggaacc	5040
ccgacgtgca	gcagatcctg	cgccaggaga	gcctggccgc	cgcagccagc	atcagtgaac	5100
atccccagaa	ggcaaccacc	gagctgccct	tgtgcggggc	ggccctcaag	gagaccttga	5160
ggtgggtgct	ggctgaggcc	tccctgtggc	cctggccccc	tgtggagag	cagccccac	5220
tgggtgggtg	cagacagaat	ctggggctga	taaacagcgt	caccagcag	cccatcccc	5280

tgcacctgct	cttctctccc	ctcaaggaca	gggagctctt	cttctctctgg	aatccctctt	5340
caacgccttg	gggattaacg	tggggcatgt	ccttctgcgc	tcggggctgc	ttaagttagg	5400
ggaggttttg	ctgggctcag	caggtgcaag	gaagcacttc	ctacgacctg	ggcttcccat	5460
ggatctggga	cctctgcggg	gtcttcggta	ggaaggggtg	tgagagcaca	gggagcccca	5520
tccagctgag	gacctttct	gtggatgccc	ccacctccag	gctctaccct	gtgggtctgt	5580
ttttggagcg	agtggtagag	tcagacttgg	tgcttcagaa	ctaccacatc	ccagctgggg	5640
tgagttagcc	ccacaccctc	cgagctgaga	acctccctcc	ccagtcattc	cctgatccct	5700
gctctgcacc	gtccgcagac	attggtacag	gttttctctt	actcgctggg	tcgcaatgcc	5760
gccttggtcc	cgaggcctga	gcggtataat	ccccagcgct	ggctagacat	caggggctcc	5820
ggcaggaact	tccaccacgt	gccctttggc	tttggcatgc	gccagtgcct	cgggcggcgc	5880
ctggcagagg	cagagatgct	gctgctgctg	caccacgtaa	gcaggcctgg	gggcgggggc	5940
gggacctggg	cagcagaggc	gggacctgca	cactgggggc	ggggcttgca	tggtgtgatt	6000
gacacctggg	aacagtggat	ggggccttgg	ttggttgagg	tcggcgtgac	cagggaggat	6060
ctgtgctgag	caagacaggg	taggatctgg	gtgaggttgc	ttctaaacat	tgaaatgggg	6120
actaggggag	tggggtggag	cctgtacaga	ataatggggc	ttgggcaaga	cctgggcagg	6180
attcagtcctg	ggcctggtcc	gcaaggtggg	gctggtcaga	aatgggatag	gttggggccc	6240
aggctgctgc	tcccccttca	gcataattgt	tgcacctggg	acgatgggag	gaagctgccc	6300
cagggtccatg	ggctactgac	caggccagat	ggaaacccag	cctctgtcct	aggtgctgaa	6360
gcacttctctg	gtggagacac	taactcaaga	ggacataaag	atggtctaca	gcttcatatt	6420
gaggcctggc	acgtccccc	tcctcacttt	cagagcgatt	aactagtctt	gcatctgcac	6480
ccagggtccc	agcctggcca	ccagcttccc	tctgcctgac	cccaggccac	ctgtcttctc	6540
tcccacgtgc	acagcttccc	gagtcacccc	tctgtccagc	cagctcctgc	acaaatggaa	6600
ctccccaggg	cctccaggac	tggggcttgc	caggcttgtc	aaatagcaag	gccagggcac	6660
agctggagac	gatcttgctg	gcagggcctg	gccttgtccc	cagccccacc	tgcccccttc	6720
tccagcaagc	agtgcctctt	ggacagcttg	actctactcc	tcccagcgct	ggctccaggc	6780
tcctcatgag	gccatgcaag	gggtgctgtg	ttttgtccct	tgccttctctg	cctagtctca	6840
catgtccctg	tccctctcgc	cctggccagg	gcctctgtgc	agacagtgtc	agagtcatta	6900
agcgggatcc						6910

<210> 135

<211> 2476

<212> DNA

<213> Homo sapien

<400> 135

gaattcttgt	tttacaagcc	atggctctgt	ttcttaattgt	tttctataat	cactcacttt	60
ttttttgctt	ttgacaaaca	ttcaaaatgc	taatgattca	aggatgtcct	cagctctgta	120
tgtgttctaa	gagttctatg	ttttttctcc	acagaaggca	taagaactag	gagctgtgta	180
catttcaata	tgaagggcaa	ctccaccctt	gccactacta	gcaaaaacat	taccagcggt	240
cttcaacttcg	ggcttgtgaa	catctctggc	aacaatgagt	ctaccttgaa	ctgttcacag	300
aaaccatcag	ataagcattt	agatgcaatt	cctattcttt	actacattat	atttgttaatt	360
ggatttctgg	tcaatatgtt	cgtgggttaca	ctgttttgtt	gtcaaaaggg	tcctaaaaag	420
gtttctagca	tatacatctt	caacctcgct	gtggctgatt	tactcctttt	ggctactctt	480
cctctatggg	caacctatta	ttcttataga	tatgactggc	tctttggacc	tgtgatgtgc	540
aaagtttttg	gttcttttct	tacctgaac	atgtttgcaa	gcattttttt	tatcacctgc	600
atgagtgttg	ataggtacca	atctgtcatc	tacctcttcc	tgtctcaaag	aagaaatccc	660
tggcaagcat	cttatatagt	tccccttggt	tgggtgatgg	cctgtttgtc	ctcattgcca	720
acatttttatt	ttcgagacgt	cagaaccatt	gaatacttag	gagtgaatgc	ttgcattatg	780
gctttccacc	ctgagaaata	tgcccaatgg	tcagctggga	ttgccttaat	gaaaaatatc	840
cttgggtttta	ttatcccttt	aatattcata	gcaacatgct	attttggaat	tagaaaaaac	900
ttactgaaga	cgaatagcta	tgggaagaac	aggataaccc	gtgaccaagt	cctgaagatg	960
gcagctgctg	ttgttctggc	cttcatcatt	tgtctggctc	ccttccatgt	tctgaccttc	1020
ctggatgctc	tggcctggat	gggtgtcatt	aatagctgcy	aagttatagc	agtcattgac	1080

ctggcacttc	cttttgccat	cctcttggga	ttcaccaaca	gctgcgttaa	tcggtttctg	1140
tattggtttg	ttggaaaccg	gttccaacag	aagctccgca	gtgtgttttag	ggttccaatt	1200
acttggtctc	aagggaanaag	agagagtatg	tcttgccgga	aaagcagttc	tcttagagaa	1260
atggagacct	ttgtgtctta	aacgtgagag	caaaatgcat	gtaatcaaca	tggtactttg	1320
ctttgaggct	caccagaatt	atttttaagt	gggttttaata	aaataataaa	atttccccta	1380
atcttttctg	aatcttctga	aaccaaatgt	aactatgttt	tatcgtccag	tgactttcag	1440
gaattgcccc	ttgtttttct	gatattgtttg	tacaagattt	tcattgggtga	gacatattta	1500
caacctagaa	gtaactgggtg	atatactca	aattgtaatt	aataatagat	tgtgaataat	1560
gatttgggga	ttcagatttc	tctttgaaac	atgcttgtgt	ttcttagtgg	ggttttatat	1620
ccatttttat	caggattttcc	tcttgaacca	gaaccagtct	ttcaactcat	tgcatcattt	1680
acaagacaac	attgtaagag	agatgagcac	ttctaagtgtg	agtatattat	aatagattag	1740
tactggatta	ttcaggcttt	aggcatatgc	ttctttaaaa	acgctataaa	ttatattcct	1800
cttgcatctc	acttgagtgg	aggtttatag	ttaatctata	actacatatt	gaatagggtc	1860
aggaatatag	attaaatcat	actcctatgc	tttagcttat	ttttacagtt	atagaaagca	1920
agatgtacta	taacatagaa	ttgcaatcta	taatatattgt	gtgttcacta	aactctgaat	1980
aagcactttt	taaaaaactt	tctactcatt	ttaatgattg	tttaaagggt	tctattttct	2040
ctgatacttt	tttgaatca	gtaaacactg	tgtattgttg	taaaatgtaa	aggtcacttt	2100
tcacatcctt	gacttttttag	atgtgctgct	ttgatataata	ggacattgat	ttgattttta	2160
ttactaatgc	tttggttctg	ggttggttcc	taaaatatct	gggtggccta	aaaaaaactc	2220
tttaacttgt	aataaaacct	taactggcat	aggaaatggg	atccagaatg	gaattttgct	2280
acatgggggtc	tgggtggggg	caaagagacc	cagtcaatta	catgtttggg	accaagaaaag	2340
gaacctgtca	gggcagtaca	atgtgacttt	gaaaatatat	accgtggggg	tagttttacc	2400
ctatatctat	aaacactgtt	tgttccagaa	tctgtatgat	tctatggagc	tatttttaaac	2460
caattgcagg	tctaga					2476

<210> 136

<211> 3100

<212> DNA

<213> Homo sapien

<400> 136

gaattcctga	gatccctgta	ctggtaactcc	tgctagaata	accttctggt	ttcgagggtt	60
tataatttaa	tgtattatac	agaaattcgg	agaaacttga	tgtatctctg	aatgagcaga	120
gaaaaaaagt	ttcaggaaag	ttttccaaca	ttccaatgtc	acgtgtgttt	atatatatac	180
tgtgtacggc	aaagttaactg	gcaatctggc	ttctgccctg	aacctctacg	ccaagttctg	240
gccttatcga	cacagaaaag	caatgccttc	cacccttcgg	gggcatttaa	ggttgacaca	300
ctctattccg	gtaacaccac	ttgacagaaa	gatggcaagg	aaaagcctct	tccccacatt	360
ccctatgcct	caccctgcag	ctcagcactc	agaactcctt	tcctgggcag	cgtaatttc	420
aacttcttcc	aatcccttgg	ctatttttag	actgccccct	gcgggtccag	aagggaagcca	480
ggaggccggg	gcagtggctc	atgcctctat	ccagcacttt	gggaggccaa	ggtgggtgga	540
tcacctgagc	ttaggagttc	gagacgagcc	taccaacatg	gtaaaacgcc	atctccacta	600
aaaatacaaa	attagccgtg	tgtggcggca	cacgcctgta	atcccagcta	cttgacaggc	660
tgaggcagga	gaatcacttg	aacccgggag	gccagagggt	gcagtgggcc	aagatcacgc	720
actgaactcc	agcctgggca	acaaaagtga	aactctgtct	caaaaaata	aaaagaaaaga	780
cgccaggagt	ctcaggcttt	cactgacttg	gataacttgg	ggtgtagtct	ccttctctcc	840
ctttctaaag	ctcttgggtt	cctggccact	cctcctaacc	cccccaagtt	ggaaaaatcag	900
ctgataaatc	atgggcaggg	ggtgaataaa	caaatcactc	cccagtttta	catactacat	960
atgatgctga	ggtttggggc	taaaagcaaat	tcatatgttt	gcattttatgt	aagagaaagt	1020
ggttaaaact	ggttgatatt	atataaaact	gaaaaacatt	actttttgct	tttgacctgg	1080
acagctctgg	aagtcggtct	ctccttgcaa	ggtaaggata	tgagcttggg	catattgatg	1140
accatcaacg	aacatgatgt	tcgttacggg	cgtgtttctt	attccatgct	agaaggacca	1200
gtgagcatta	tataacattt	tcattgcacc	ttacataaac	aattagaagc	tgtttactcc	1260
tcaaggactc	aattagagat	tttcttaatt	tttttatttt	tttttttatt	tcaggccctga	1320
gctgaggaca	ctacaggagc	caaattagag	aaggggggct	gcagtccctg	catattgcac	1380
attcatggct	ttcttgtccc	ctcaccccc	gtataatcct	ttttgccagg	tttaacaactc	1440

actaatattca	aagcaattca	tttgccaact	caaaaaaaaa	aaaaaaaaaa	aaaaccacaca	1500
gttccaactt	taaattgctt	ttggattttta	atcctctggg	gtattttatat	gtttccttgg	1560
ctcatccctc	caaaactaat	ctcagagttt	tgggaatctg	ggaacttggg	caaaggggga	1620
aaccaggcac	gaagaccaag	gatctgaaat	cgcagttcat	tacgtcagca	taaaactgaca	1680
gtaccttttt	gtttgtattg	acccagctca	aattataaaa	tcacatgagt	aataaaacac	1740
aaaatacagg	tgttatcatt	gaaaagtcta	aaatgtaatt	aaacgcgttt	ttccccctcg	1800
cgggtggtacg	tttactctgt	atctcaacgt	ttctgtctcc	ctaagcctct	cctttcatat	1860
catgagcata	cattttttgca	tcattgctcac	acgttcatta	gctaggactg	aggagtgtgt	1920
acgattccaa	agggccctca	gcgttagctg	ttagatgcac	aaaccttcgc	ttcctttcca	1980
catctactgc	actcgaggtt	caacagagga	tcttgaagaa	cagcgccaga	gagcattctt	2040
gacagatgcg	cgcttggttc	cagcaaccgc	tctgctggga	agtttcttct	aaccactaac	2100
accacctcca	atccccaag	ctgtcacgac	gcatgctggc	tgggtcccgt	cttgacgggg	2160
gaaggggtttt	acacacaccc	tgctaggctg	ccccacatca	caaccaagct	cgcagggcaa	2220
actcctccaa	gcctggcgga	caagctgtcc	caggcgctct	ggcgcttcct	gaacaccaag	2280
gtccccctccc	ccgtcaaggg	agctagcgct	ctgttcgcga	gagaaccccg	gaactgcagg	2340
tcgaggggat	gcggggggag	cgggggcgca	ggaggagcc	gagtgcctgga	ggcaaacggg	2400
gcgcaggagc	gggtgcggga	ggcaaacggg	gcgcaggagt	gggtgcggga	ggcaaacggg	2460
cgcaggagtgc	gggtgcgggag	cagagtggggg	ctcaggagggt	gtcgggcccgc	ggggagccag	2520
gcgcggaagg	gggcgcgggg	gaacagggac	caggaaccag	cgggcgcagg	aagggtgctg	2580
tcggaaggaa	cccgcgggcg	cacgggaggc	actagctacg	cgatcagctc	gggactctca	2640
ggagccgctc	aattgccaac	gggagggggg	tcgggggagt	tggaggtggg	ggagcagacc	2700
agacgggggc	gtgcctttgc	ccgcattggc	tgaggagacc	tgacgcgagg	ccccgggggt	2760
tggcttgggg	agtgggagcg	gggtgggggtg	ggtgctgggt	gccggagctg	cgggcccggc	2820
gcgctcagaa	acatgctgaa	gtcccggcgg	ctcttcagac	agcaggcagc	ggctccagca	2880
gcagcaggca	ggcaggcagg	caggcaggca	ggcagcggca	gcgacagcgc	tcggctctgc	2940
gggaaaacgg	gcccgggcgc	catgctccgg	ccccgcgggg	cggctgcctc	gacccggccg	3000
cgacctccct	ctgcgcacca	cgcgcgcgg	gcttctgggg	tggtcccca	ccacggccca	3060
gccttgcac	acccccgcg	cccggcctcc	gcagctcggc			3100

<210> 137

<211> 1723

<212> DNA

<213> Homo sapien

<400> 137

tgctaccgcg	gcccgggctt	ctgggggtgt	ccccaaaccac	ggcccagccc	tgccacaccc	60
cccgcceccg	gcctccgcag	ctcggcattg	gcgcgggggt	gctcgtcctg	ggcgccctcg	120
agcccggtaa	cctgtcgtcg	gcgcacccgc	tcccgcagcg	cgcggccacc	gcggcgccgc	180
tgctgggtgc	cgcgtcgccg	cccgcctcgt	tgctgcctcc	cgccagcgaa	agcccgcagc	240
cgtctgtctca	gcagtggaca	gcgggcatgg	gtctgctgat	ggcgtcctc	gtgctgtctca	300
tcgtggcggg	caatgtgctg	gtgatcgtgg	ccatcgccaa	gacgccgcgg	ctgcagacgc	360
tcaccaacct	cttcatcatg	tccttgccca	gcgcgcacct	ggtcatgggg	ctgctggtgg	420
tgccgttcgg	ggccaccatc	gtgggtgtgg	gccgctggga	gtacggctcc	ttcttctgcg	480
agctgtggac	ctcagtggac	gtgctgtgcg	tgacggccag	catcgagacc	ctgtgtgtca	540
ttgcccctgga	ccgctacctc	gccatcacct	cgcctctccg	ctaccagagc	ctgctgacgc	600
gcgcgcgggc	gcggggcctc	gtgtgcaccg	tgtgggccat	ctcgccctcg	gtgtccttcc	660
tgcccatcct	catgcactgg	tgggggggcg	agagcgacga	ggcgccgcgc	tgctacaacg	720
accccaagtg	ctgcgacttc	gtcaccaccc	ggccttacgc	catcgccctcg	tcogtagtct	780
ccttctacgt	gcccctgtgc	atcatggcct	tcgtgtacct	gcgggtgttc	cgcgaggccc	840
agaagcaggt	gaagaagatc	gacagctcgc	agcgcggttt	cctcgccggc	ccagcgcggc	900
cgccctcgcc	ctcgccctcg	cccgtcccgc	cgcgcgcgcg	gccgcgccga	ccccgcgcgc	960
ccgcgcgcgc	cgccgcacc	gccccgctgg	ccaacggggc	tgccgggtta	cggcgccctc	1020
cgccgctcgt	ggccctacgc	gagcagaagg	cgctcaagac	gctgggcatc	atcatgggcg	1080
tcttcacgct	ctgctggctg	cccttcttcc	tgcccaacgt	ggtgaaggcc	ttccaccgcg	1140
agctgggtgc	cgaccgcctc	ttcgtcttct	tcaactggct	gggctacgcc	aactcggcct	1200

tcaaccccat	catctactgc	cgcagccccc	acttccgcaa	ggccttccag	ggactgctct	1260
gctgcgcgcg	cagggctgcc	cgcggcgccc	acgcgaccca	cggagaccgg	ccgcgcgcct	1320
cgggctgtct	ggcccggccc	ggaccccccgc	catcgcccgg	ggccgcctcg	gacgacgacg	1380
acgacgatgt	cgtcggggcc	acgccgcccc	cgcgcctgct	ggagccctgg	gccggctgca	1440
acggcggggc	ggcggcggac	agcgactcga	gcctggacga	gccgtgccgc	cccggcttcg	1500
cctcggaatc	caaggtgtag	ggcccggcgc	ggggcgcgga	ctccgggcac	ggcttcccag	1560
gggaacgagg	agatctgtgt	ttactttaaga	ccgatagcag	gtgaactcga	agcccacaat	1620
cctcgtctga	atcatccgag	gcaaagagaa	aagccacgga	ccgttgcaaa	aaaaggaaaag	1680
tttgggaagg	gatgggagag	tggttgcgtg	atgttccttg	ttg		1723

<210> 138

<211> 3451

<212> DNA

<213> Homo sapien

<400> 138

cccggttca	agagattctc	ctgtctcagc	ctccccagta	gctgggacta	caggtacgtg	60
ccaccacacc	tggctaattt	ttgtattttt	agtagagaca	agagttacac	catattggcc	120
aggatctttt	gctttctata	gcttcaaaat	gttcttaatg	ttaagacatt	cttaatactc	180
tgaaccatat	gaatttgcca	ttttggtaag	tcacagacgc	cagatggtgg	caatttcaca	240
tggcacaacc	cgaagatta	acaaactatc	cagcagatga	aaggattttt	tttagtttca	300
ttgggtttac	tgaagaaatt	gtttgaattc	tcattgcac	tccagttcaa	cagataatga	360
gtgagtgatg	ccacactctc	aagagttaaa	aacaaaacaa	caaaaaaatt	aaaacaaaag	420
cacacaactt	tctctctctg	tcccaaaata	catacttgca	tacccccgct	ccagataaaa	480
tccaaagggt	aaaactgtct	tcattgcctgc	aaattcctaa	ggaggggacc	taaaagtactt	540
gacagcgagt	gtgctgagga	aatcggcagc	tggtgaagtc	acctcctgtg	ctcttgccaa	600
atgtttgaaa	gggaatacac	tgggttaccg	ggtgtatgtt	gggaggggag	cattatcagt	660
gctcgggtga	ggcaagtctg	gagtaccacg	atggagacat	ccgtgtctgt	gtcgtctctg	720
atgcctccaa	gccagcgtgt	gtttactttc	tggtgtgtgc	accatgtctt	tggtctctctg	780
gggtctctctg	tggttctctg	tggcgcgctt	tctgtgttgg	acaggggtga	ctttgtgccg	840
gatggcttct	gtgtgagagc	gcgcgcgagt	gtgcatgtcg	gtgagctggg	aggggtgtgtc	900
tcagtgtcta	tggctgtggt	tcggtataag	tctgagcatg	tctgccaggg	tgtatttgtg	960
cctgtatgtg	cgtgcctcgg	tgggcactct	cgtttccttc	cgaatgtggg	gcagtgcggg	1020
tggtcgtccc	tctgccttga	gacctcaagc	cgcgcaggcg	cccagggcag	gcaggtagcg	1080
gccacagaag	agccaaaagc	tcccgggttg	gctggtaagg	acaccacctc	cagcttttagc	1140
cctctggggc	cagccagggt	agccgggaag	cagtgggtggc	ccgccctcca	gggagcagtt	1200
gggccccgcc	cgggccagcc	ccaggagaag	gagggcgagg	ggaggggagg	gaaaggggag	1260
gagtgccctg	ccccctcggg	gctgccggcg	tgccattggc	cgaaggttcc	cgtacgtcac	1320
ggcgagggca	gttcccceta	agtcctgtgc	acataacggg	cagaacgcac	tgcaagcggg	1380
cttcttcaga	gcacgggctg	gaactggcag	gcaccgcgag	cccctagcac	ccgacaagct	1440
gagtggtcag	gacgagtccc	caccacaccc	acaccacagc	cgctgaatga	ggcttccagg	1500
cgtccgctcg	cggccccgag	agccccgcgc	tgggtccgcc	cgctgaggcg	ccccagcca	1560
gtgcgcttac	ctgccagact	gcgcgccatg	gggcaaccgc	ggaacggcag	cgcttctctg	1620
ctggcaccaca	atagaagcca	tgccgcggac	cacgacgtca	cgcagcaaag	ggacgaggtg	1680
tgggtggtgg	gcatgggcat	cgtcatgtct	ctcatcgtcc	tggccatcgt	gtttggcaat	1740
gtgctggtca	tcacagccat	tgccaagttc	gagcgtctgc	agacgggtcac	caactacttc	1800
atcacttcac	tggcctgtgc	tgatctggtc	atgggccttg	cagtgggtgcc	ctttggggcc	1860
gcccatattc	ttatgaaaat	ttggactttt	ggcctaactct	ggtgcgagtt	ttggactttc	1920
attgatgtgc	tgtgcgtcac	ggccagcatt	gagaccctgt	gcgtgatcgc	agtggatcgc	1980
tacttttgcca	ttacttcacc	tttcaagtac	cagagcctgc	tgaccaagaa	taaggccccg	2040
gtgatcattc	tgatggtgtg	cattgtgtca	ggccttacct	ccttcttgcc	cattcagatg	2100
cactgggtacc	gggccaccca	ccaggaagcc	atcaactgct	atgccaatga	gacctgctgt	2160
gacttcttca	cgaaccaagc	ctatgccatt	gcctcttcca	tcgtgtcctt	ctacgttccc	2220
ctggtgatca	tggctctcgt	ctactccagg	gtcttttcagg	aggccaaaag	gcagctccag	2280
aagattgaca	aatctgaggg	ccgcttccat	gtccagaacc	ttagccaggt	ggagcaggat	2340

gggaggacgg	ggcatggact	ccgcagatct	tccaagttct	gcttgaagga	gcacaaagcc	2400
ctcaagacgt	taggcatcat	catgggcact	ttcaccctct	gctggctgcc	cttcttcatc	2460
gttaacattg	tgcattgtat	ccaggataac	ctcatccgta	aggaagttaa	catctccta	2520
aattggatag	gctatgtcaa	ttctggtttc	aatcccccta	tctactgccg	gagcccagat	2580
ttcaggattg	ccttccagga	gcttctgtgc	ctgcgcagg	cttctttgaa	ggcctatggg	2640
aatggctact	ccagcaacgg	caacacaggg	gagcagagt	gatatacagt	ggaacaggag	2700
aaagaaaata	aactgctgtg	tgaagacctc	ccaggcacgg	aagactttgt	gggccatcaa	2760
ggtactgtgc	ctagcgataa	cattgattca	caaggaggga	attgtagtac	aaatgactca	2820
ctgctgtaaa	gcagtttttc	tacttttaaa	gacccccccc	cccccaacag	aacactaaac	2880
agactattta	acttgagggt	aataaaactta	gaataaaatt	gtaaaaattg	tatagagata	2940
tgcagaagga	agggcatcct	tctgcctttt	ttattttttt	aagctgtaaa	aagagagaaa	3000
acttatttga	gtgattattt	gttatttcta	cagttcagtt	cctctttgca	tggattttgt	3060
aagtattatgt	ctaaagagct	ttagtcctag	aggacctgag	tctgctatat	tttcatgact	3120
tttccatgta	tctacctcac	tattcaagta	ttaggggtaa	tattattgctg	ctggttaattt	3180
gtatctgaag	gagattttcc	ttcctacacc	cttggacttg	aggattttga	gtatctcgga	3240
cctttcagct	gtgaacatgg	actcttcccc	cactcctctt	atttgctcac	acggggtatt	3300
ttaggcaggg	atttgaggag	cagcttcagt	tgttttcccg	agcaaagggtc	taaagtttac	3360
agtaaaataa	atgtttgacc	atgccttcat	tgcacctgtt	tgtccaaaac	cccttgactg	3420
gagtgtctgt	gcctccccc	ctggaaaccg	c			3451

<210> 139

<211> 2305

<212> DNA

<213> Homo sapien

<400> 139

gaattcatgc	cgcttttctg	tgttggacag	gggtgacttt	gtgccggatg	gcttctgtgt	60
gagagcgccg	gcgagtgtgc	atgtcgggtga	gctgggaggg	tgtgtctcag	tgtctatggc	120
tgtggttcgg	tataagtcta	agcatgtctg	ccagggtgta	tttgtgcctg	tatgtgcgtg	180
cctcgggtggg	cactctcggt	tccttccgaa	tgtggggcag	tgccgggtgtg	ctgcctctg	240
ccttgagacc	tcaagccggc	cagggcccca	gggcaggcag	gtagcggcca	cagaagagcc	300
aaaagctccc	gggttggctg	gtaagcacac	cacctccagc	tttagccctc	tggggccagc	360
cagggtagcc	gggaagcagt	gggtggcccg	cctccaggga	gcagttgggc	cccggccggg	420
ccagcctcag	gagaaggagg	gcgaggggag	gggagggaaa	ggggaggagt	gcctcgcccc	480
ttcgcggctg	ccggcggtgc	attggccgaa	agttcccgta	cgtcacggcg	agggcagttc	540
ccctaaagtc	ctgtgcacat	aacgggcaga	acgcactgcg	aagcggcttc	ttcagagcac	600
gggctggaac	tggcaggcac	cgcgagcccc	tagcacccca	caagctgagt	gtgcaggacg	660
agtcgccacc	acacccacac	cacagccgct	gaatgaggct	tccaggcgctc	cgctcgccgg	720
ccgcagagcc	ccgcccgtgg	tccgcctgct	gaggcgcccc	cagccagtgc	gcttacctgc	780
cagactgcgc	gccatggggc	aacccgggaa	cggcagcgcc	ttcttgctgg	cacccaatag	840
aagccatgcg	ccggaccacg	acgtcacgca	gcaaaaggag	gaggtgtggg	tgggtgggcat	900
gggcatcgtc	atgtctctca	tcgtcctggc	catcggtgtt	ggcaatgtgc	tgggtcatcac	960
agccattgcc	aagtctcgagc	gtctgcagac	ggtcaccaac	tacttcatca	cttcaactgg	1020
ctgtgtctgt	ctgggtcatgg	gcctggcagt	gggtgccctt	ggggccgccc	atattcttat	1080
gaaaatgtgg	acttttggca	acttctgggtg	cgagttttgg	acttccattg	atgtgctgtg	1140
cgtcacggcc	agcattgaga	ccctgtgcgt	gatcgagtg	gatcgctact	ttgccattac	1200
ttcacctttc	aagtaccaga	gcctgctgac	caagaataag	gcccgggtga	tcatttctgat	1260
gggtgtggatt	gtgtcaggcc	ttacctcctt	cttgcccatt	cagatgcact	ggtaccgggc	1320
cacccaccag	gaagccatca	actgctatgc	caatgagacc	tgctgtgact	tcttcacgaa	1380
ccaagccctat	gccattgcct	cttccatcgt	gtccttctac	gttcccctgg	tgatcatggg	1440
cttcgtctac	tccaggggtc	ttcaggaggc	caaaaggcag	ctccagaaga	ttgacaaatc	1500
tgagggccgc	ttccatgtcc	agaaccttag	ccaggtggag	caggatgggc	ggacggggca	1560
tggactccgc	agatcttcca	agttctgctt	gaaggagcac	aaagccctca	agacgttagg	1620
catcatcatg	ggcactttca	ccctctgctg	gctgcccctc	ttcatcgtaa	acattgtgca	1680
tgtgatccag	gataacctca	tccgtaagga	agttttacatc	ctcctaattt	ggataggcta	1740

tgtcaattct	ggtttcaatc	cccttatcta	ctgccggagc	ccagatttca	ggattgcctt	1800
ccaggagctt	ctgtgcctgc	gcagggtctc	tttgaaggcc	tatgggaatg	gctactccag	1860
caacggcaac	acagggggagc	agagtggata	tcacgtggaa	caggagaaaag	aaaataaaact	1920
gctgtgtgaa	gacctcccag	gcacgggaaga	ctttgtgggc	catcaaggta	ctgtgcctag	1980
cgataacatt	gattcacaag	ggaggaattg	tagtacaaat	gactcactgc	tgtaaagcag	2040
tttttctact	tttaaagacc	cccccccca	acagaacact	aaacagacta	tttaacttga	2100
gggtaataaa	cttagaataa	aattgtaaaa	ttgtatagag	atatgcagaa	ggaagggcat	2160
ccttctgcct	tttttatttt	tttaagctgt	aaaaagagag	aaaacttatt	tgagtgatta	2220
tttggtattt	gtacagttca	gttcctcttt	gcatggaatt	tgtaagttta	tgtctaaaga	2280
gctttagtc	tagaggacct	gagtc				2305

<210> 140

<211> 4105

<212> DNA

<213> Homo sapien

<400> 140

gaattcgcgg	ccgcctcttg	cgggtcccaga	gtggagtgga	aggctctggag	ctttggggag	60
agacggggag	gacagactgg	aggcgtgttc	ctccggagtt	ttctttttcg	tcgagagccct	120
cgcgcgcgcg	tacagtcctc	ccgctgggtc	gacgattgtg	gagaggcggg	ggagaggcctt	180
catccatccc	acccgggtcgt	cgcgggggat	tgggggtccca	gcgacacctc	cccggggagaa	240
gcagtgccca	ggaagttttc	tgaagccggg	gaagctgtgc	agccgaagcc	gccgcgcgcg	300
cggagcccg	gacaccggcc	accctccgcg	ccaccacccc	tcgctttctc	cggcttctctc	360
tggcccaggc	gccgcgcggga	cccgccagct	gtctgcgcac	gccgagctcc	acggtgaaaa	420
aaaaagtga	gggtgtaaaag	cagcacaagt	gcaataagag	atatttctct	aaatttgctt	480
caagatggaa	accctttgcc	tcagggcatc	cttttggctg	gcactggttg	gatgtgtaat	540
cagtataat	cctgagagat	acagcacaaa	tctaagcaat	catgtggatg	atttcaccac	600
ttttcgtggc	acagagctca	gcttcctggt	taccactcat	caaccacta	atttgggtcct	660
acccagcaat	ggctcaatgc	acaactattg	cccacagcag	actaaaatta	cttcagcttt	720
caaatacatt	aacactgtga	tatcttgtac	ttttttcatc	gtgggaatgg	tggggaatgc	780
aactctgctc	aggatcattt	accagaacaa	atgtatgagg	aatggcccca	acgcgctgat	840
agccagtcct	gcccttggag	accttatcta	tgtgggtcatt	gatctcccta	tcaatgtatt	900
taagctgctg	gctgggcgct	ggccttttga	tcacaatgac	tttggcgat	ttctttgcaa	960
gctgttcccc	tttttgcaga	agtccctcgt	ggggatcacc	gtcctcaacc	tctgcgctct	1020
tagtggtgac	aggtacagag	cagttgcctc	ctggagtcgt	gttcagggaa	ttgggattcc	1080
tttggttaact	gccattgaaa	ttgtctccat	ctggatcctg	tcctttatcc	tggccattcc	1140
tgaagcgatt	ggcttcgtca	tggtaaccct	tgaatatagg	gggtgaacagc	ataaaacctg	1200
tatgctcaat	gccacatcaa	aattcatgga	gttctaccaa	gatgtaaaag	actggtggct	1260
cttcgggttc	tatttctgta	tgcccttggt	gtgcaactgc	atcttctaca	ccctcatgac	1320
ttgtgagatg	ttgaacagaa	nnaatggcag	cttgagaatt	gccctcagtg	aacatcttaa	1380
gcagcgtcga	gaagtggcaa	aaacagtttt	ctgcttgggt	gtaatttttg	ctctttgctg	1440
gttcctctct	cacttaagcc	gtatatggaa	gaaaactgtg	tataacgaaa	tggacaagaa	1500
ccgatgtgaa	ttacttagtt	tcttactgct	catggattac	atcggtatta	acttggcaac	1560
catgaattca	tgtataaacc	ccatagctct	gtattttgtg	agcaagaaat	ttaaaaattg	1620
tttcagtcga	tgctctgct	gctgctgtta	ccagtcctaaa	agtctgatga	cctcgggtccc	1680
catgaacgga	acaagcatcc	agtgggaaga	ccacgatcaa	aacaaccaca	acacagaccg	1740
gagcagccat	aaggacagca	tgaactgacc	acccttagaa	gcactcctcg	gtactcccat	1800
aatcctctcg	gagaaaaaaa	tcacaaggca	actgtgactc	cggaatctc	ttctctgatc	1860
cttcttctct	aattcactcc	cacacccaag	aagaaatgct	ttccaaaaacc	gcaaggtaga	1920
ctgggtttatc	caccacaac	atctacgaat	cgtacttctt	taattgatct	aatttacata	1980
ttctgcgtgt	tgtattcagc	actaaaaaat	ggtgggagct	gggggagaat	gaagactgtt	2040
aaatgaaacc	agaaggatat	ttactacttt	tgcataaaaa	tagagctttc	aagtacatgg	2100
ctagctttta	tggcagttct	ggtgaatggt	caatgggaac	tggtcaccat	gaaacttttag	2160
agattaacga	caagattttc	tacttttttt	agtgtatttt	ttgtccttca	gccaaacaca	2220
atatgggctc	aggtcacttt	tatttgaaat	gtcatttgggt	gccagtatct	tttaactgca	2280

taatagccta	acatgattat	ttgaacttat	ttacacatag	tttgaaaaaa	aaaagacaaa	2340
aatagtatcc	aggtgagcaa	ttagattagt	attttccacg	tcactattta	tttttttaaa	2400
acacaaattc	taaagctaca	acaaatacta	caggccctta	aagcacagtc	tgatgacaca	2460
tttggcagtt	taatagatgt	tactcaaaga	attttttaag	aactgtattt	tattttttaa	2520
atgggtgttt	attacaaggg	accttgaaca	tgttttgtat	gttaaattca	aaagtaatgc	2580
ttcaatcaga	tagttctttt	tcacaagttc	aatactgttt	ttcatgtaaa	ttttgtatga	2640
aaaatcaatg	tcaagtacca	aaatgttaat	gtatgtgtca	tttaactctg	cctgagactt	2700
tcagtgcact	gtatatagaa	gtctaaaaa	cacctaaagag	aaaaagatcg	aatttttcag	2760
atgattcgga	aattttcatt	caggtatttg	taatagtga	atataatgt	atatacatat	2820
cacctcctat	tctcttaatt	tttgttaaaa	tgtaactgg	cagtaagtct	tttttgatca	2880
ttcccttttc	catataggaa	acataatttt	gaagtggcca	gatgagttta	tcagtgcagt	2940
gaaaaataat	taccacaaaa	tgccaccagt	aacttaacga	ttcttcactt	cttgggggtt	3000
tcagtatgaa	cctaactccc	cacccaaca	tctccctccc	acattgtcac	catttcaaag	3060
ggccacagtc	gacttttgct	gggcattttc	ccagatgttt	acagactgtg	agtacagcag	3120
aaaatccttt	actagtgtgt	gtgtgtatat	atataaaca	ttgtaaat	cttttagccc	3180
atttttctag	actgtctctg	tggaatatat	ttgtgtgtgt	gatataatga	tggtgtgat	3240
ggatatgtatg	gatttaattct	aatctaataa	ttgtgccccg	cagttgtgcc	aaagtgcata	3300
gtctgagcta	aaatctaggt	gattgttcat	catgacaacc	tgccctcagtc	cattttaacc	3360
tgtagcaacc	ttctgcattc	ataaatcttg	taatcatgtt	accattacaa	atgggatata	3420
agaggcagcg	tgaaagcaga	tgagctgtgg	actagcaata	taggggtttg	tttggtttgt	3480
tggtttgata	aagcagtatt	tggtgtcata	ttgtttcctg	tgctggagca	aaagtcatta	3540
cactttgaag	tatttatattg	ttcttatcct	caattcaatg	tggtgatgaa	attgccaggt	3600
tgctgatat	ttctttcaga	cttcgccaga	cagattgtcg	ataataaatt	aggtaagata	3660
atttgttggg	ccatatattta	ggacaggtaa	aataacatca	ggttccagtt	gcttgaattg	3720
caaggctaag	aagtactgcc	cttttgtgtg	ttagcagtc	aatctattat	tcactggcg	3780
catcatatgc	agtgatatat	gcctataata	taagccatag	gttcacacca	ttttgtttag	3840
acaattgtct	ttttttcaag	atgctttgtt	tctttcatat	gaaaaaaatg	cattttataa	3900
attcagaaag	tcataagattt	ctgaaggcgt	caacgtgcat	tttatttatg	gactggtaag	3960
taactgtggt	ttactagcag	gaatatttcc	aatttctacc	tttactacat	cttttcaaca	4020
agtaactttg	tagaaatgag	ccagaagcca	aggccctgag	ttggcagtg	cccataagtg	4080
taaaaataaaa	gtttacagaa	acctt				4105